

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 15:34:28 Search time 457.59 Seconds
(without alignments)
2352.554 Million cell updates/sec

Title: US-09-824-647-16_COPY_1_627

Perfect score: 627

Sequence: 1 cgcaggcagaccattggac.....cagtgccctgtccagctcg 627

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_032802.*

1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 627 | 100.0 | 2095 | 20 AAV82825 | Human GP88 autocri |
| 2 | 627 | 100.0 | 2124 | 14 AAQ56794 | Granulin coding se |
| 3 | 627 | 100.0 | 2198 | 14 AAQ49052 | Granulin coding se |
| 4 | 627 | 100.0 | 2432 | 21 AAC78180 | Human cancer assoc |
| 5 | 615 | 98.1 | 1779 | 12 AAQ14339 | Human epithelin pr |
| 6 | 406.6 | 64.8 | 1767 | 12 AAQ14338 | Rat epithelin prec |
| 7 | 400.8 | 63.9 | 2137 | 20 AAV82824 | Mouse GP88 autocri |
| 8 | 390.4 | 62.3 | 1767 | 12 AAQ14340 | Mouse epithelin pr |
| 9 | 381.2 | 60.8 | 561 | 21 AAA69780 | Human ovarian carc |

| | | | | | |
|----|-------|------|---------|-------------|---------------------|
| 10 | 297.4 | 47.4 | 757 | 22 AAH07141 | Human cDNA clone (|
| 11 | 297.4 | 47.4 | 1630 | 22 AAH16370 | Human cDNA sequenc |
| 12 | 59 | 9.4 | 341 | 12 AAQ14953 | Chicken epithelin |
| 13 | 50.2 | 8.0 | 51 | 21 AAC15693 | Human secreted pro |
| 14 | 48.6 | 7.8 | 539 | 12 AAQ14952 | Bovine epithelin p |
| 15 | 41.6 | 6.6 | 3729 | 23 AAS94337 | DNA encoding novel |
| 16 | 40.2 | 6.4 | 10732 | 21 AAQ10594 | Gene encoding a su |
| 17 | 40 | 6.4 | 1824 | 23 AAS81488 | DNA encoding novel |
| 18 | 40 | 6.4 | 2243 | 22 ABA08657 | Human 'extensin hom |
| 19 | 40 | 6.4 | 2772 | 23 AAS72787 | DNA encoding novel |
| 20 | 40 | 6.4 | 2772 | 23 AAS90525 | DNA encoding novel |
| 21 | 40 | 6.4 | 2850 | 23 AAS79895 | DNA encoding novel |
| 22 | 39.2 | 6.3 | 799 | 19 AAV55831 | Nucleotide sequenc |
| 23 | 39.2 | 6.3 | 1925 | 20 AAX90924 | Epstein Barr Virus |
| 24 | 39.2 | 6.3 | 1926 | 21 AAQ50254 | Epstein Barr virus |
| 25 | 39.2 | 6.3 | 1926 | 22 AAF82902 | EBV tethering prot |
| 26 | 39.2 | 6.3 | 2580 | 21 AAQ75454 | Nucleotide sequenc |
| 27 | 39.2 | 6.3 | 5452 | 20 AAX90923 | Anti-sense strand |
| 28 | 39.2 | 6.3 | 8705 | 20 AAQ23778 | Vector pShuttle DN |
| 29 | 39.2 | 6.3 | 9600 | 19 AAV21683 | Nucleotide sequenc |
| 30 | 39.2 | 6.3 | 10380 | 20 AAQ22248 | Plasmid pCisEBON f |
| 31 | 39.2 | 6.3 | 10596 | 14 AAQ51731 | Plasmid pCisEBON f |
| 32 | 39.2 | 6.3 | 10596 | 17 AAQ40348 | Nucleotide sequenc |
| 33 | 39.2 | 6.3 | 10596 | 20 AAQ15850 | DNA clone pCEK Cl. |
| 34 | 39.2 | 6.3 | 16080 | 21 AAS95953 | DNA encoding novel |
| 35 | 38.4 | 6.1 | 109519 | 22 AAS08693 | DNA encoding novel |
| 36 | 38 | 6.1 | 1267 | 23 AAS67932 | DNA encoding novel |
| 37 | 38 | 6.1 | 1833 | 23 AAS67931 | Murine transport a |
| 38 | 37.8 | 6.0 | 565 | 22 AAK53492 | Lysine decarboxyla |
| 39 | 37 | 5.9 | 3269 | 17 AAT34583 | Human brain expres |
| 40 | 36.8 | 5.9 | 233 | 22 AAK18681 | Aspergillus niger |
| 41 | 36.8 | 5.9 | 2028 | 15 AAQ73737 | Aspergillus niger |
| 42 | 36.8 | 5.9 | 2476 | 15 AAQ73736 | Myobacterium tube |
| 43 | 36.6 | 5.8 | 4403765 | 22 AAI99683 | Myobacterium tube |
| 44 | 36.6 | 5.8 | 4411529 | 22 AAI99682 | P. putida KT2440-a |
| 45 | 36.2 | 5.8 | 1680 | 22 AAF61077 | |

ALIGNMENTS

RESULT 1

AAV82825

ID AAV82825 standard; cDNA; 2095 BP.

XX AAV82825;

AC AAV82825;

XX 15-MAR-1999 (first entry)

XX Human GP88 autocrine growth factor cDNA.

XX GP88; granulin; epithelin; human; growth factor; autocrine; tumour;
XX cancer; viral infection; antagonist; therapy; diagnosis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 13..1794

XX /*tag= a

XX WO9852607-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-US10555.

XX 16-DEC-1997; 97US-0991862.

XX 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

XX Serrero G;

XX

| | |
|----|---|
| KW | allergic reaction; graft versus host disease; organ rejection; |
| KW | haemostatic; thrombolytic; cardiovascular disorder; infection; |
| KW | neurological disease; drug screening; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| XX | WO200055350-A1. |
| PN | |
| PD | 21-SEP-2000. |
| XX | |
| PF | 08-MAR-2000; 2000WO-US05882. |
| XX | |
| XX | 12-MAR-1999; 99US-0124270. |
| PR | |
| XX | (HUMA-) HUMAN GENOME SCI INC. |
| XX | |
| PI | Rosen CA, Ruben SM; |
| XX | |
| XX | WPI; 2000-587533/35. |
| DR | P-PSDB; AAB43971. |
| DR | |
| XX | |
| PT | Novel isolated nucleic acids comprising sequences encoding peptides |
| PT | useful for treating or diagnosing e.g. cancer - |
| XX | |
| PS | Claim 1; Page 1097-1098; 2352pp; English; |
| XX | |
| XX | AAC77607 to AAC78448 encode the human cancer associated proteins given |
| CC | in ABA43398 to ABA44239. The proteins can have activities based on the |
| CC | tissues and cells the genes are expressed in. Example of activities |
| CC | include: cytostatic; proliferative; vulnery; immunomodulator; |
| CC | antidiabetic; antiasthmatic; antirheumatic; antiarthritic; |
| CC | antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral; |
| CC | dermatological; neuroprotective; cardiant; thrombolytic; coagulant; |
| CC | neotropic; vasotropic; antipsoriatic and antiangiogenic. The |
| CC | polynucleotides and polypeptides can be used for preventing, treating or |
| CC | ameliorating medical conditions and diagnosing pathological conditions. |
| CC | Polynucleotides, polypeptides, antibodies, agonists and antagonists from |
| CC | the present invention may be used to treat immune disorders by activating |
| CC | or inhibiting the proliferation, differentiation or mobilisation of |
| CC | immune cells, to treat disorders of haematopoietic cells, autoimmune |
| CC | disorders, allergic reactions, graft versus host disease and organ |
| CC | rejection, modulate haemostatic or thrombolytic activity, modulate |
| CC | inflammation, cancers, cardiovascular disorders, neurological disease and |
| CC | bacterial or viral infections. The peptides, nucleotides, antibodies, |
| CC | agonists and antagonists may be also be used in drug screens. AAC78449 to |
| CC | AAC78457 and ABA4240 represent sequences used in the exemplification of |
| CC | the present invention. |
| XX | |
| XX | Sequence 2432 BP; 460 A; 764 C; 707 G; 496 T; 5 other; |

[illegible]

| | | | |
|----------|----------|---|-----|
| Db | 416 | gtctcagggaattccgaattgctgccccttcccagaggccgctggcatgcggggatggcccat | 475 |
| QY | 301 | cactgtgcccacggggtctccactgcagtcagacagggcgatcctgttccaaagatca | 360 |
| Db | 476 | cactgtgcccacggggtctccactgcagtcagacagggcgatcctgttccaaagatca | 535 |
| QY | 361 | ggtacaaactcgggtgggtgccatccagtcgacctgatagtcagttcgaatgccggacttc | 420 |
| Db | 536 | ggtaacaaactcgggtgggtgccatccagtcgacctgatagtcagttcgaatgccggacttc | 595 |
| QY | 421 | ttcacatgtgttatgtatgcattgctcctgaggggtgctgcccacatgcccgacttc | 480 |
| Db | 596 | ttcacatgtgtgttatgtatgcattgctcctgaggggtgctgcccacatgcccgacttc | 655 |
| QY | 481 | tgctgtgaagacagggtgcactgctgcgcacggctgcttctgcgacctggttcacacc | 540 |
| Db | 656 | tgctgtgaagacagggtgcactgctgcgcacggctgcttctgcgacctggttcacacc | 715 |
| QY | 541 | cgctgcatcacacccacgggacaccccccctggcaagaagctcctgcccagaggact | 600 |
| Db | 716 | cgctgcatcacacccacgggacaccccccctggcaagaagctcctgcccagaggact | 775 |
| QY | 601 | aacagggcagtgcttgcctccagctcg | 627 |
| Db | 776 | aacagggcagtgcttgcctccagctcg | 802 |
| RESULT 5 | | | |
| QY | AAQ14339 | AAQ14339 standard; DNA; 1779 BP. | |
| XX | XX | AAQ14339; | |
| XX | XX | 17-JAN-1992 (first entry) | |
| XX | XX | Human epithelin precursor. | |
| XX | XX | ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss. | |
| XX | XX | Homo sapiens. | |
| XX | XX | Key Location/Qualifiers | |
| XX | XX | misc_RNA 41..1819 | |
| XX | XX | /*tag= a | |
| XX | XX | /note= "claim 9, page 54" | |
| XX | XX | WO9115510-A. | |
| XX | XX | 17-OCT-1991. | |
| XX | XX | 03-APR-1991; 91WO-US02321. | |
| XX | XX | 13-MAR-1991; 91US-0083796. | |
| XX | XX | 03-APR-1990; 90US-0504508. | |
| XX | XX | (BRIM) BRISTOL-MYERS SQUIB. | |
| XX | XX | Shoyab M, Plowman GD; | |
| XX | XX | WPI; 1991-325168/44. | |
| XX | XX | P-PSDB; AARI4336. | |
| XX | XX | New cysteine-rich growth modulating proteins, epithelins - useful | |
| XX | XX | as inhibitors of neoplastic cell growth and to promote wound | |
| XX | XX | healing and treat psoriasis | |
| XX | XX | Disclosure; Fig 22; 97pp; English. | |
| XX | XX | ET-1 and ET-2 were isolated from rat kidneys and their amino acid | |
| XX | XX | sequences determined. A full length rat ET CDNA (AAQ14338) was obtnd. | |
| XX | XX | by screening a rat kidney cDNA library in lambda gt10 with PCR | |
| XX | XX | generated ET probes. These probes were also used to obtain the mouse | |
| XX | XX | ET gene (AAQ14340) from a mouse T-cell genomic library. | |

ET DNA was also obtained from human sources (AAQ14339). An anti-sense ribonucleic acid molecule complementary to the indicated fragment in the features is also claimed. ET-1 is a bifunctional growth regulator, capable of stimulating the growth of some cell types while inhibiting the growth of others. ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory bioactivity. In contrast, however, ET-2 is apparently not capable of eliciting the growth stimulatory activity characteristic of ET-1 and, in fact, antagonises this ET-1 activity. See also AAR14338-40, AAR14952-53, AAR14328-9 and AAR15315-20.

```
Query Match      98.1%; Score 615; DB 12; Length 1779;
Best Local Similarity 100.0%; Pred. No. 2.1e-151;
Matches 615: Conservative 0; Mismatches 1 0; Indels 0; Gaps 0;
```

| | | | |
|----|-----|--|-----|
| QY | 13 | atgtgacctggtgagctgggtgaccttaacagcagggctggtggcctggaaacgcgggtgc | 72 |
| Db | 1 | atgtgacctggtgagctgggtgaccttaacagcagggctggtggcctggaaacgcgggtgc | 60 |
| QY | 73 | ccagatgggtcagtttgcctgtgacctgtgacctggaaccccgaggagccagctacagc | 132 |
| Db | 61 | ccagatgggtcagtttgcctgtgacctgtgacctggaaccccgaggagccagctacagc | 120 |
| QY | 133 | tgtgcgcgtccctcttgacaaaatggcccaaacactgagcagcactctgggtggcccc | 192 |
| Db | 121 | tgtgcgtccctcttgacaaaatggcccaaacactgagcagcactctgggtggcccc | 180 |
| QY | 193 | tgcaggttgatgcccaactgctctgcggccactcctgcatctttaccgctcagggact | 252 |
| Db | 181 | tgcaggttgatgcccaactgctctgcggccactcctgcatctttaccgctcagggact | 240 |
| QY | 253 | tccagttgtgcaccttccacagaggcgtgggaatgcggggatggccactcactgtctgccca | 312 |
| Db | 241 | tccagttgtgcaccttccacagaggcgtgggaatgcggggatggccactcactgtctgccca | 300 |
| QY | 313 | cggaggttccaactgcagtcagtcagacggcgcatcctgcttccaaagatcaggttaacaactcc | 372 |
| Db | 301 | cggaggttccaactgcagtcagtcagacggcgcatcctgcttccaaagatcaggttaacaactcc | 360 |
| QY | 373 | gtgggtgcacatccagtgccctgatatgctcagtttggaaatgcccggaacttctccacgtgctgt | 432 |
| Db | 361 | gtgggtgcacatccagtgccctgatatgctcagtttggaaatgcccggaacttctccacgtgctgt | 420 |
| QY | 433 | gttatggtcgatggctcctgggggtgctgcccaatgcccaaggcttctctgctgtgaagac | 492 |
| Db | 421 | gttatggtcgatggctcctgggggtgctgcccaatgcccaaggcttctctgctgtgaagac | 480 |
| QY | 493 | agagtgacgtgcttcgcgacggtgcttctgcgacctggttcacaccgcgtgcacacaa | 552 |
| Db | 481 | agagtgacgtgcttcgcgacggtgcttctgcgacctggttcacaccgcgtgcacacaa | 540 |
| QY | 553 | ccacagggcacccaccccttgcacaaagaagctccctgccagaggactaacaggcagtg | 612 |
| Db | 541 | ccacagagacccaccccttgcacaaagaagctccctgccagaggactaacaggcagtg | 600 |

| | |
|----------|-----------------------------------|
| RESULT | 6 |
| AAQ14338 | |
| ID | AAQ14338 standard; cDNA; 1767 bp. |
| XX | |
| XX | AAQ14338; |
| XX | |
| DT | 17-JAN-1992 (first entry) |
| XX | |
| DE | Rat epithelin precursor. |
| XX | |

ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.

Rattus rattus.

| Key | Location/Qualifiers |
|----------|----------------------------|
| misc_RNA | 31..1797 |
| | /*tag= a |
| | /note= *claim 19, page 55" |

WO9115510-A.

17-OCT-1991.

03-APR-1991; 91WO-US02321.

13-MAR-1991; 91US-0083796.

03-APR-1990; 90US-0504508.

(BRIM) BRISTOL-MYERS SQUIB.

Shoyab M, Plowman GD;

WPI: 1991-325168/44.

P-PSDB; AARI4325.

New cysteine-rich growth modulating proteins, epithelins - useful as inhibitors of neoplastic cell growth and to promote wound healing and treat psoriasis

Disclosure; Fig 18; 97pp; English.

ET-1 and ET-2 were isolated from rat kidneys and their amino acid sequences determined. A full length rat ET cDNA was obtd. by screening a rat kidney cDNA library in lambda gt10 with PCR generated ET probes. These probes were also used to obtain the mouse ET gene (AAQ14340) from a mouse T-cell genomic library. ET DNA was also obtained from human sources (AAQ14339)...

An anti-sense ribonucleic acid molecule complementary to the indicated fragment in the features is also claimed.

ET-1 is a bifunctional growth regulator, capable of stimulating the growth of some cell types while inhibiting the growth of others.

ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory bioactivity. In contrast, however, ET-2 is apparently not capable of eliciting the growth stimulatory activity characteristic of ET-1 and, in fact, antagonises this ET-1 activity.

See also AAQ14338-40, AAQ14952-53, AARI4328-9 and AARI5315-20.

Sequence 1767 BP: 359 A: 499 C: 491 G: 418 T: 0 other;

| | | | | |
|---------------------------|--------|--------------------|-----------|--------------|
| Query Match | 64.8%; | Score 406.6; | DB 12; | Length 1767; |
| Best Local Similarity | 80.7%; | Pred. No. 7.7e-97; | | |
| Matches 488: Conservative | 0; | Mismatches 114; | Indels 3; | Gaps 1; |

| | | | |
|----|-----|--|-----|
| QY | 13 | atgtgacacctggtgagctgggttggcccttaacagcaggagctggtgagtcggaacgcggtgc | 72 |
| | | | |
| | | | |
| Db | 1 | atgtggaacctggtgagctggctttagtggaagctggtggtggaacacagtcg | 60 |
| | | | |
| QY | 73 | ccagatggtcagttctgcctctggtgaccccgaggagccagctacagc | 132 |
| | | | |
| | | | |
| Db | 61 | ccagaagggtcaattctgacctgtgtcctgctctgacaggaggagcccaactacagc | 120 |
| | | | |
| QY | 133 | tgtcgcgtccccctcttggaacaatggccacaacactgagcaggcatctggttggcccc | 192 |
| | | | |
| | | | |
| Db | 121 | tgtgtaaacctctctggacacatggcctataatacagagcgtcgctagtaggtccc | 180 |
| | | | |
| QY | 193 | tgcagggtgatgccactgtctcgcgccactctctgcatcttaccgtctcagggaact | 252 |
| | | | |
| Db | 181 | tgcagatccggtgaccactgtcctctgattggctactctgtcttctcaactgtgtctgggaact | 240 |
| | | | |
| QY | 253 | tccagttgctgcccttccacagaggcgtggcatgcggggatggccatcactgctgtgccca | 312 |
| | | | |
| | | | |
| Db | 241 | tcaagatctcccattctctgaagatctatcttgtatgatggcagacactgctgtccc | 300 |
| | | | |

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 XX EP1074617-A2.
 XX 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-0116126.
 XX
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-0118776.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 XX full-length cDNAs defined in the specification, and for the detection
 XX and/or diagnosis of the abnormality of the proteins encoded by the
 XX full-length cDNAs.
 XX
 XX Claim 8; SEQ ID 15309; 2537pp + CD ROM; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602
 XX full-length cDNAs defined in the specification. Where a primer set
 XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 XX to the complementary strand of a polynucleotide which comprises one of
 XX the 5602 nucleotide sequences defined in the specification, where the
 XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 XX of an oligonucleotide comprising a sequence complementary to the
 XX complementary strand of a polynucleotide which comprises a 5'-end
 XX sequence and an oligonucleotide comprising a sequence complementary to a
 XX polynucleotide which comprises a 3'-end sequence, where the
 XX oligonucleotide comprises at least 15 nucleotides and the combination of
 XX the 5'-end sequence/3'-end sequence is selected from those defined in
 XX the specification. The primer sets can be used in antisense therapy and
 XX in gene therapy. The primers are useful for synthesizing polynucleotides,
 XX particularly full-length cDNAs. The primers are also useful for the
 XX detection and/or diagnosis of the abnormality of the proteins encoded by
 XX the full-length cDNAs. The primers allow obtaining of the full-length
 XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 XX represent oligonucleotides, all of which are used in the exemplification
 XX of the present invention.
 XX
 XX Sequence 1630 BP; 285 A; 517 C; 489 G; 339 T; 0 other;

Query Match 47.4%; Score 297.4; DB 22; Length 1630;
 Best Local Similarity 99.7%; Pred. No. 2.9e-68;
 Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 209 actgtctgcggccactctgcatctttaccgtctcagggaacttcagggtgtgcccct 268
 DB 1 actgtctgcggccactctgcatctttaccgtctcagggaacttcagggtgtgcccct 60
 QY 269 tccagagggccgtggcgtcggggatggccatcactgtcccacgggggttccactgca 328
 DB 61 tccagagggccgtggcgtcggggatggccatcactgtcccacgggggttccactgca 120
 QY 329 gtgcagagggccgtatcctgtttccaaagatcaaggtaaacactcgtgggtgccatcaggt 388
 DB 121 gtgcagagggccgtatcctgtttccaaagatcaaggtaaacactcgtgggtgccatcaggt 180
 QY 389 gccctgatgtcattcgaatgcccggaacttctccacgtgtgtgttatgggtcgtggtcct 448

DB 181 gcctgatagtcagttcgaatgccggaacttctccacgtgtgtgttatgtcgtgct 240
 QY 449 cctgggggtgctgccccatgccaggtctcctgctgtagaacagaggtgctgctgct 507
 DB 241 cctgggggtgctgccccatgccaggtctcctgctgtagaacagaggtgctgctgct 299
 RESULT 12
 AAQ14953
 ID AAQ14953 standard; DNA; 341 BP.
 XX
 XX AC AAQ14953;
 XX
 XX 17-JAN-1992 (first entry)
 XX
 XX Chicken epithelin precursor (partial).
 DE
 XX KW ET; growth regulation; inhibition; stimulation; ss.
 XX OS Gallus domesticus.
 XX PN W09115510-A.
 XX PD 17-OCT-1991.
 XX PF 03-APR-1991; 91WO-US02321.
 XX PR 13-MAR-1991; 91US-0083796.
 XX PS 03-APR-1990; 90US-0504508.
 XX (BRIM) BRISTOL-MYERS SQUIB.
 XX Shoyab M; Plowman GD;
 WPI; 1991-325168/44.
 P-PSDB; AARI5427.
 New cysteine-rich growth modulating proteins, epithelins - useful
 as inhibitors of neoplastic cell growth and to promote wound
 healing and treat psoriasis
 Disclosure; Fig 25; 97pp; English.
 The epithelins appear to comprise several distinct members sharing
 significant structural homology. Two members of the epithelin family,
 EP-1 and EP-2, have been purified from natural sources, and cDNAs
 encoding these and several other members of the epithelin family have
 been isolated from rat (AAQ14338), human (AAQ14339), bovine (AAQ14952),
 murine (AAQ14340) and chicken (AAQ14953).
 EP-1 is a bifunctional growth regulator, capable of stimulating
 the growth of some cell types while inhibiting the growth of others.
 EP-2 is functionally similar to EP-1 w.r.t. growth inhibitory
 bioactivity. In contrast, however, EP-2 is apparently not capable of
 eliciting the growth stimulatory activity characteristic of EP-1 and,
 in fact, antagonises this EP-1 activity.
 See also AAQ14338-40, AAQ14952-53, AARI4328-9 and AARI5315-20.
 Sequence 341 BP; 53 A; 108 C; 120 G; 59 T; 1 other;

Query Match 9.4%; Score 59; DB 12; Length 341;
 Best Local Similarity 62.2%; Pred. No. 4.9e-06;
 Matches 92; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
 QY 451 tgggggtgctgccccatgccaggtctcctgctgtagaacagaggtgctgctgctc 510
 DB 4 tgggggtgctgccccatgccaggtgctgctgtagaacagaggtgctgctgctc 63
 QY 511 caggtgcttctgctgccccatgccaggtgctgtagaacagaggtgctgctgctc 570
 DB 64 cactccaccagctgtgattggagcgcggcgctgtgtgtccctacggggaagctccc 123

[illegible]

Search completed: September 5, 2002, 18:30:01

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|------------------|-------------------|
| 1 | 627 | 100.0 | 2095 | 4 | US-08-991-862-16 | Sequence 16, Appl |
| 2 | 615 | 98.1 | 1779 | 1 | US-07-668-648-3 | Sequence 3, Appl |
| 3 | 615 | 98.1 | 1779 | 2 | US-08-429-998-3 | Sequence 3, Appl |
| 4 | 615 | 98.1 | 1779 | 2 | US-08-431-333-3 | Sequence 3, Appl |
| 5 | 615 | 98.1 | 1779 | 5 | PCT-US91-02321-3 | Sequence 3, Appl |
| 6 | 406.6 | 64.8 | 1767 | 1 | US-07-668-648-1 | Sequence 1, Appl |
| 7 | 406.6 | 64.8 | 1767 | 2 | US-08-429-998-1 | Sequence 1, Appl |
| 8 | 406.6 | 64.8 | 1767 | 2 | US-08-431-333-1 | Sequence 1, Appl |
| 9 | 406.6 | 64.8 | 1767 | 5 | PCT-US91-02321-1 | Sequence 1, Appl |
| 10 | 399.2 | 63.7 | 2137 | 4 | US-08-991-862-1 | Sequence 1, Appl |
| 11 | 390.4 | 62.3 | 1767 | 1 | US-07-668-648-5 | Sequence 5, Appl |
| 12 | 390.4 | 62.3 | 1767 | 2 | US-08-429-998-5 | Sequence 5, Appl |
| 13 | 390.4 | 62.3 | 1767 | 2 | US-08-431-333-5 | Sequence 5, Appl |
| 14 | 390.4 | 62.3 | 1767 | 5 | PCT-US91-02321-5 | Sequence 5, Appl |
| 15 | 59 | 9.4 | 341 | 5 | PCT-US91-02321-9 | Sequence 9, Appl |
| 16 | 58.4 | 9.3 | 341 | 1 | US-07-668-648-9 | Sequence 9, Appl |
| 17 | 58.4 | 9.3 | 341 | 2 | US-08-429-998-9 | Sequence 9, Appl |
| 18 | 58.4 | 9.3 | 341 | 2 | US-08-431-333-9 | Sequence 9, Appl |
| 19 | 48.6 | 7.8 | 539 | 1 | US-07-668-648-7 | Sequence 7, Appl |
| 20 | 48.6 | 7.8 | 539 | 2 | US-08-429-998-7 | Sequence 7, Appl |
| 21 | 48.6 | 7.8 | 539 | 2 | US-08-431-333-7 | Sequence 7, Appl |
| 22 | 48.6 | 7.8 | 539 | 5 | PCT-US91-02321-7 | Sequence 7, Appl |
| 23 | 39.2 | 6.3 | 2580 | 3 | US-09-050-863-2 | Sequence 2, Appl |
| 24 | 39.2 | 6.3 | 2580 | 4 | US-09-359-081-2 | Sequence 2, Appl |
| 25 | 39.2 | 6.3 | 5452 | 2 | US-09-130-114-1 | Sequence 1, Appl |
| 26 | 39.2 | 6.3 | 9600 | 1 | US-08-910-647-1 | Sequence 1, Appl |
| 27 | 39.2 | 6.3 | 10596 | 1 | US-07-884-811-15 | Sequence 15, Appl |


```

QY      613 gcctgtccagctcg 627
        |||||
Db      601 GCCTGTCCAGCTCG 615

RESULT      4
US-08-431-333-3
; Sequence 3, Application US/08431333
; Patent No. 5965723
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25

```

```

APPLICATION NUMBER: 087874317533
FILING DATE: 27-APR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/668,648
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1779 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1779
US-08-431-333-3

Query Match          98.1%; Score 615; DB 2; Length 1779;
Best Local Similarity 100.0%; Pred. No. 1.3e-153;
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 13 atgtggacctggtgagctgggtggccttaacagcagggtcgtggtcggaaacgcgggtgc 72
    |||
Db 1 ATGTGGACCCCTGGGTGAGCTGGGTGGCCCTTAACAGCAGGGCTGGTGGCTGGAAACGGGGTGC 60

QY 73 ccagatggtcagttctgcctctgtggcctctgcctcggaccccgaggagcgacgtcacagc 132
    |||
Db 61 CCAGATGGTCTAGTCTGTGCCCTGTGGCTGTGCTTGGACCCCGGAGGAGCCAGCTACAGC 120

QY 133 tgcctgcgcgtcccccctctctggacaaatggccccacaacactgagcaggcatctcgtgggtggcccc 192
    |||
Db 121 TGCYGGCCCTCCCTCTCTGGCAAAATGGCCCAACACACTGAGCAGGSCATCTGGGTGGCCGCC 180

```

193 tgcaggttgatgccactgtctgcccggccactctcatctttaccgtctcaggact 252
181 tccaggttgatgccactgtctgcccggccactctcatctttaccgtctcaggact 240
253 tccaggttgatgccactgtctgcccggccactctcatctttaccgtctcaggact 312
241 tccaggttgatgccactgtctgcccggccactctcatctttaccgtctcaggact 300
313 cggggttcctccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 372
301 cggggttcctccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 360
373 gtgggtgcatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 432
361 gtgggtgcatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 420
433 gttatggtcgatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 492
421 gttatggtcgatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 480
493 aggggtgcatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 552
481 aggggtgcatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 540
553 cccagggcaccaccccccctgagtcagagcggcgtctcttccaaagatcaggttaacactcc 612
541 cccagggcaccaccccccctgagtcagagcggcgtctcttccaaagatcaggttaacactcc 600
613 gcttgcagctcg 627
601 gcttgcagctcg 615

RESULT 5
PCT-US91-02321-3
; Sequence 3, Application PC/TUS9102321
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plovman, Gregory D.
; TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02321
; FILING DATE: 19910403
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: ON0071A-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)728-4800
; TELEFAX: (206)448-4775
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1779 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1779
PCT-US91-02321-3

Query Match 98.1%; Score 615; DB 5; Length 1779;
Best Local Similarity 100.0%; Pred. No. 1.3e-153;
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 13 atgtggacctggtgagctgggtgaccttaacagcagggctggtgctggaacagcgggtgc 72
Db 1 ATGTGGACCTGGTGGCTGGCTTAAACAGCAGGGCTGGTGGCTGGTGGCTGGTGGCTGG 60
Qy 73 ccagatgactgcttgcctgctgctgctgctgctgctgctgctgctgctgctgctgctg 132
Db 61 CCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 133 tgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 192
Db 121 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy 193 tgcaggttgatgcccactgctctgcccggccactctcttaccgtctcaggact 252
Db 181 TGCCAGGTTGATGCCACTGCTCTGCGGCCACTCTCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy 253 tccagttgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 312
Db 241 TCCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy 313 cggggttcctccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 372
Db 301 CGGGGTTCCTCCACTGAGTCAGAGCGGGCGATCTGCTTCCAAAGATCAGGTAACTCC 360
Qy 373 gtgggtgcatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 432
Db 361 GTGGGTGCATCCACTGAGTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 433 gttatggtcgatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 492
Db 421 GTTATGTCGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 493 aggggtgcatccactgagtcagagcggcgtctcttccaaagatcaggttaacactcc 552
Db 481 AGGGTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 553 cccagggcaccaccccccctgagtcagagcggcgtctcttccaaagatcaggttaacactcc 612
Db 541 CCCAGGGCACCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 613 gcttgcagctcg 627
Db 601 GCTTGTCCAGCTCG 615

RESULT 6
US-07-668-648-1
; Sequence 1, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plovman, Gregory D.
; TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

```

Db 538 CCCACGGGCCCCACCCCTTACTTAAGAAATTCCTCCCGCAAAAGGACCAACAGGGCAGTG 597
QY 613 gcctt 617
Db 598 GCCTT 602

RESULT 7
US-08-429-998-1
: Sequence 1, Application US/08429998
: Patent No. 585961
: GENERAL INFORMATION:
: APPLICANT: Shoyab, Mohammed
: APPLICANT: Plowman, Gregory D.
: TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
: TITLE OF INVENTION: MODULATING PROTEINS
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

? FILING DATE: 27-APR-1995
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/668,648
? FILING DATE: 13-MAR-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Misrock, S. Leslie
? REGISTRATION NUMBER: 18,872
? REFERENCE/DOCKET NUMBER: 5624-161-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212)790-9090
? TELEFAX: (212) 869-9741
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1767 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? HYPOTHETICAL: NO
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1767
? US-08-429-998-1

? Query Match 64.8%; Score 406.6; DB 2; Length 1767;
? Best Local Similarity 80.7%; Pred. No. 1.4e-98;
? Matches 488; Conservative
?
? Qy 13 atgtggacctgtgagctgggtgacctaaacagcagggctggtgctggaacgcggtgc 72
? ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
? Db 1 ATGTGGATCTCTGGTGACCTGGCTGGCCCTAGTGTGCGAAGCGCTGCTGGCTGGACACAGTCG 60
?
? Qy 73 ccagatggtcagttctgccctgtggcctgtgctgctggaaccccgaggagccagctacagc 132
? ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
? Db 61 CCAGATGGTCAATTCTGCCCTGTGTGCTGTGCTGTGCTGTGACCGAGGAGGCAACTACAGC 120
?
? Qy 133 tgcgtgcgctccctcttcggacaataaggcccaaacactgagcagcgatctgagtgcccc 192
? ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
? Db 121 TCGTGTGAACCTCTTCTCGGACATGGCTTAAATACAGAGCGCGTCTAGATGCTCC 180

```

| | | | | |
|---------------------------|--------|--------------------|-----------|--------------|
| Query Match | 64.8%; | Score 405.6; | DB 2; | Length 1767; |
| Best Local Similarity | 80.7%; | Pred. No. 1.4e-98; | | |
| Matches 488: Conservative | 0; | Mismatches 114; | Indels 3; | Gaps 1; |

```

, TYPE: nucleic acid
, STRANDEDNESS: double
, TOPOLOGY: linear
, MOLECULE TYPE: CDNA
, HYPOTHETICAL: NO
, FEATURE:
, NAME/KEY: CDS
, LOCATION: 1..1767
US-08-431-333-1

```

| Query Match | 64.8%; | Score 406.6; | DB 2; | Length 1767; |
|-----------------------|-----------------|---|-----------|--------------|
| Best Local Similarity | 80.7%; | Pred. No. 1.4e-98; | | |
| Matches 488; | Conservative 0; | Mismatches 114; | Indels 3; | Gaps |
| QY | 13 | atgtgacacctggtgagctggggtggccttaacagcagggctgggtgctgggaacgggtgc 72 | | |
| DB | 1 | | | |
| QY | 73 | ccagatggttggcttctgacctgtggtgacctgacctgacctgacctgacctgacctgacctg 132 | | |
| DB | 61 | CCAGATGGTCAATTCCTGCCCTGTGCTGCTGCTTGAACAGGGAGAGCCAACTACAGC 120 | | |
| QY | 133 | tgctgcggtccctctctgacaaatggcccaaacactgacgagcagcagctggtggagcccc 192 | | |
| DB | 121 | TGCTGTAAACCTCTCTGGACACATGGCCCTATATAACGAGCCGTGCTGTAGATGGCTCC 180 | | |
| QY | 193 | tgcaggttgatgccactgctgcgggccactcctgcctctttaccgtctcagggact 252 | | |
| DB | 181 | TGCCAGATCGTGACCACGTCTGATGGCTACTCTTGTCTTCTACTGTGCTGGGACT 240 | | |
| QY | 253 | tcagttgctgaccttccagagggccgtggcatgcggggattggccatcactgctgtccca 312 | | |
| DB | 241 | TCCAGCTGCTGCCCGTCTCTGAGGGTGATCTTGTGATGATGGCCAGGACACTGCTGCCCC 300 | | |
| QY | 313 | cggggcttccactgcagtcgagcgggcatcctgcttccaaagatcaggtaaacactcc 372 | | |
| DB | 301 | CGGGGCTTCCACTGTAGTCGGATGGGAACTCTGCTCTCAGATATCAGATAGCCCTCT -- 358 | | |
| QY | 373 | gtgggtgcaatcaagtgcacctgatatgactgattgcaatgcccggacttctccagtgctgt 432 | | |
| DB | 359 | -TGGGTGCTGTCCAGTGTCTCTGTGTGCCAGTTCGAAATGCTTGACTCCGCCACCTGCTGT 417 | | |
| QY | 433 | gttatggtcgatggctcctgggggtgctgccccatgcccgacttctctgctgtgaagac 492 | | |
| DB | 418 | ATTATGATTGATGTTCTCTGGGGGTGCTGCCCATGCCCCAGGCCCTCTTGTGCTGAAGAC 477 | | |
| QY | 493 | agggtgcaactgctgcgcgaacgggtgcttctgcgaacctggttcacaccgcgtgcatcaca 552 | | |
| DB | 478 | AGAGTGCATTGCTGTCCCAACGGGGGCTCTCTGTGACTGCTGTTACACGGCGATGCATTTCA 537 | | |
| QY | 553 | ccacagggcacccaccctctggcaagaagcttccctgcccagaggactaacagggcgctg 612 | | |
| DB | 538 | CCACAGGGCACCCACCCCTTACTTAAGAAATTCCTCCGCGCAAAAGGACCAACAGGGCAGTG 597 | | |
| QY | 613 | gcctt 617 | | |
| DB | 598 | GCCTT 602 | | |

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/668,648
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs

STATE: Washington
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02321
FILING DATE: 19910403
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: POOR, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE/DOCKET NUMBER: ON0071A-PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)728-4800
TELEFAX: (206)448-4775
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Rat
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1767
PCT-US91-02321-1

Db 478 AGAGTGCATTGCTGCCACGGGGCTCTGTGACCTGTTGCACAGCGATCATTTCA 537
QY 553 CCACAGGCGACCCACCCCTGGCAAGAGCTCCCTGCCAGAGGACTAACAGGCGAGTG 612
Db 538 CCACAGGCGACCCACCCCTTACTAAGAAATTCCTCCGACAAAGGACACAGGCGAGTG 597
QY 613 GCCTT 617
Db 598 GCCTT 602

RESULT 10
US-08-991-862-1
Sequence 1, Application US/08991862
Patent No. 6309826
GENERAL INFORMATION:
APPLICANT: Serrero, Ginette
TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
FILE REFERENCE: Z9996.488/P001-A
CURRENT APPLICATION NUMBER: US/08/991,862
CURRENT FILING DATE: 1998-08-17
EARLIER APPLICATION NUMBER: 08/863,862
EARLIER FILING DATE: 1997-05-23
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2137
TYPE: DNA
ORGANISM: Mouse epithelin/granulin
FEATURE:
NAME/KEY: CDS
LOCATION: (23)...(1789)
OTHER INFORMATION: The sequence is identical to that of the published
OTHER INFORMATION: mouse granulin except for one nucleotide (T
OTHER INFORMATION: instead of G) at position 1071 of GP88 cDNA
OTHER INFORMATION: (position 1056 of mouse granulin).
US-08-991-862-1

Query Match 64.8%; Score 406.6; DB: 5; Length 1767;
Best Local Similarity 80.7%; Pred. No. 1.4e-98;
Matches 488; Conservative 0; Mismatches 114; Indels 3; Gaps 1;

QY 13 atgtggacctgtgagctgggtggccttaacagcagggctgggtggcgaacggtgc 72
Db 1 ATGTGGATCTCTGTGGACATATGCTGATGAGGCTGTGGTGGTGGACACATGC 60
QY 73 ccaatagtcagctgcctgtgctgctgctgacccggagagcagcactacagc 132
Db 61 CCAGATGGTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 133 tgctgctgctccctcttggacaatggcccaacactgagcagcactctgggtggcccc 192
Db 121 TGCTGTAACCTCTCTGGACATATGCTGATATATACGAGCGCTGCTGATGCTCC 180
QY 193 tgcaggtgatgccactgctctgcggccactcctgcactcttaccgtctcaggact 252
Db 181 TGCCAGATCCGTCACCACTGTCTGTGATGGCTACTCTTGTCTCTCACTGTGTGGGACT 240
QY 253 tceagtgctgccccttccagagccgctggcagcgggagatgcccactcactgccc 312
Db 241 TCCAGCTGCTGCCGCTCTCTGAGGGGTATCTGTGATGATGGCCAGCACTGCTGCC 300
QY 313 cggggcttccactgagtcagacggcgatcctgcttccaaagatcaggttaacactcc 372
Db 301 CGGGCTTCCACTGATGGGATGGGAATCCCTGCTCTCAGATATCAGATAGCTCT-- 358
QY 373 gtgggtgccatccagtgccctgtagtagtcagttcgaaagccggagcttccacgtgtgt 432
Db 359 -TGGGTGCTGCCAGTGTCTGTGTAGGAGTTCGAATGCTCTGACTCCGCCACCTGCTGT 417
QY 433 gttatgctgatgctcctgggggtgctgcccactgcccagagcttctcgtgtgaagac 492
Db 418 ATTATGATTGATGTTCTTGGGGGTGTGCCCCATGCCAGGCTCTTGTGTGAAGAC 477
QY 493 aggtgctgctgtccgcagcgggtgctcttctgcgacctggttccacacccgctgcatcaca 552

Query Match 63.7%; Score 399.2; DB: 4; Length 2137;
Best Local Similarity 78.9%; Pred. No. 1.4e-96;
Matches 489; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

QY 1 cgacagcagaccatgtgacccctgtgagctgggtggccttaacagcagggctgtggct 60
Db 11 cgacagcagaccatgtggttctctgatgagctggctggccttcgcgcagggctgtgacc 70
QY 61 ggaacgcggtgccagatggtcagttctgacctgtggcctgtggcctgtggaccccgaggga 120
Db 71 ggaacacagtgctccagatggcagttctgacctgtgacctgtgacctgacctgacctac 130
QY 121 gcaagctacagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 180
Db 131 gcaactacagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 190
QY 181 ctgggtggccctgccaggttgatgccactgctgctgctgctgctgctgctgctgcttacc 240
Db 191 ctgagatgctcctccagaccatggcactgctgctgctgctgctgctgctgctgcttact 250
QY 241 gtctcagggacttccagttgctgctgctgctgctgctgctgctgctgctgctgctgctgct 300
Db 251 gtgctggacttccagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 310
QY 301 cactgctgcccagggcttccactgagtcagcagcagggcgagctccttcccaagatcca 360
Db 311 cactgctgcccagggcttccactgagtcagtcagtcagtcagtcagtcagtcagtcagtc 370
QY 361 ggttaacactcgtgggtgccatcagtcagtcagtcagtcagtcagtcagtcagtcagtc 420
Db 371 gataacccct---tgggtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 427
QY 421 tccacgtgctgtgttatggctgagtggtccttctgggggtgctgctgctgctgctgctgct 480

GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
MODULATING PROTEINS
TITLE OF INVENTION: MODULATING PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

AFFILIATION NUMBER: 00/00/001/000
 FILING DATE: 27-APR-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:

[illegible]

| | | | | |
|-----|--------------------|----------------------|--------------------------|-----|
| 241 | TCCAGCTGCTGCCCGTTC | TCTAAGGGTGTCTCTGTGGT | GATGGCTACCACTGCTCCCC | 300 |
| Db | | | | |
| 313 | CGGGGCTTCCACTGCAGT | GCAGACGGCGCATC | CTGTTCCAAAGATCAGGTAACAAC | 372 |
| QY | | | | |

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 13:08:22 ; Search time 3259.55 Seconds
(without alignments)
2596.245 Million cell updates/sec

Title: US-09-824-647-16_COPY_1_627

Perfect score: 627

Sequence: 1 cgcaggcagaccatgtggac.....cagtggcctgtccagctcgtc 627

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 627 | 100.0 | 653 | 10 | BF725386 |
| 2 | 627 | 100.0 | 730 | 9 | AU121637 |
| 3 | 627 | 100.0 | 760 | 9 | AU130279 |
| 4 | 627 | 100.0 | 782 | 10 | BI771292 |
| 5 | 627 | 100.0 | 785 | 10 | BG682231 |
| 6 | 627 | 100.0 | 805 | 10 | BM010556 |
| 7 | 627 | 100.0 | 814 | 9 | AU130399 |
| 8 | 627 | 100.0 | 843 | 10 | BI771009 |
| 9 | 627 | 100.0 | 849 | 9 | AU131158 |
| 10 | 627 | 100.0 | 880 | 9 | BG682258 |
| 11 | 627 | 100.0 | 920 | 9 | AL552664 |
| 12 | 627 | 100.0 | 932 | 9 | AL561424 |
| 13 | 627 | 100.0 | 986 | 9 | AL548072 |
| 14 | 627 | 100.0 | 991 | 10 | BM468251 |
| 15 | 627 | 100.0 | 1010 | 9 | AL542702 |
| 16 | 627 | 100.0 | 1106 | 10 | BM468512 |
| 17 | 627 | 100.0 | 1156 | 10 | BM478783 |

| | | | | | |
|----|-------|------|------|----|----------|
| 18 | 625.4 | 99.7 | 763 | 10 | BG824283 |
| 19 | 625.4 | 99.7 | 770 | 9 | AU132413 |
| 20 | 625.4 | 99.7 | 782 | 10 | BI755355 |
| 21 | 625.4 | 99.7 | 786 | 10 | BG831133 |
| 22 | 625.4 | 99.7 | 804 | 9 | AU125816 |
| 23 | 625.4 | 99.7 | 855 | 10 | BM009189 |
| 24 | 625.4 | 99.7 | 882 | 10 | BM007370 |
| 25 | 625.4 | 99.7 | 894 | 9 | AL557765 |
| 26 | 624.4 | 99.6 | 874 | 10 | BM051032 |
| 27 | 623.8 | 99.5 | 734 | 10 | BG823225 |
| 28 | 623.8 | 99.5 | 770 | 10 | BG469904 |
| 29 | 623.8 | 99.5 | 791 | 10 | BI918228 |
| 30 | 623.8 | 99.5 | 958 | 10 | BF312609 |
| 31 | 622.4 | 99.3 | 740 | 10 | BM045357 |
| 32 | 622 | 99.2 | 789 | 9 | AL525926 |
| 33 | 621.4 | 99.1 | 838 | 9 | AU137266 |
| 34 | 620.2 | 98.9 | 677 | 9 | AU121484 |
| 35 | 620.2 | 98.9 | 722 | 9 | AU139905 |
| 36 | 620 | 98.9 | 788 | 9 | AU124912 |
| 37 | 619 | 98.7 | 701 | 9 | AU122545 |
| 38 | 618.8 | 98.7 | 805 | 9 | AU141704 |
| 39 | 616.4 | 98.3 | 728 | 10 | BE409050 |
| 40 | 616 | 98.2 | 700 | 10 | BM046879 |
| 41 | 616 | 98.2 | 761 | 10 | BG830081 |
| 42 | 616 | 98.2 | 834 | 9 | AU122785 |
| 43 | 616 | 98.2 | 855 | 10 | BM009145 |
| 44 | 616 | 98.2 | 1033 | 9 | AL551350 |
| 45 | 615 | 98.1 | 774 | 10 | BF312922 |

ALIGNMENTS

RESULT 1
BF725386
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BF725386 653 bp mRNA linear EST 05-JAN-2001
bx15c09.y1 Human Iris cDNA (Un-normalized, unamplified): BX Homo
sapiens cDNA clone bx15c09 5', mRNA sequence.

BF725386
EST.
human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.

1 (bases 1 to 653)

NEIBANK: EST analysis and bioinformatics for ocular genomics

Invest. Ophthalmol. Vis. Sci. 41 (2000) In press

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 15 row: c column: 09

Seq primer: M13RPI reverse primer (ABI).

FEATURES

source

Location/Qualifiers

1..653

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="bx15c09"

/tissue_type="Iris"

/dev_stage="Adult"

/lab_host="EMDH10B"

/note="Organ: Eye; Vector: pcMVSPORT6; Post-mortem iris

tissue was pooled from 10 individuals ranging in age from

4-80 years and RNA was extracted. From this pooled sample

an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A

directionally cloned cDNA library in the pcMVSPORT6 vector

was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (<http://www.lifetech.com>). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGATGAGTCTAGATCGGCGGCCCTT)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing

| BASE COUNT | 106 a | 221 c | 195 g | 131 t | |
|-----------------------|--|---------------------|-----------|-------------|--|
| ORIGIN | | | | | |
| Query Match | 100.0% | Score 627; | DB 10; | Length 653; | |
| Best Local Similarity | 100.0%; | Pred. No. 8.3e-135; | | | |
| Matches 627; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; | |
| Qy 1 | cgcaggcagaccattggaacctggtgagctgggtggccttaacagcaggcgtggtggct | 60 | | | |
| Db 1 | CGCAGGCAGACCACTGTGGACCCCTGGTGAGCTGGGTGGCTTAAACAGCAGGCGTGGTGCT | 60 | | | |
| Qy 61 | ggaacgcgggtgcccaagatggtcagttctgcctgtgcccgtgctgctggaccccgaggga | 120 | | | |
| Db 61 | GGAACGGGGTGCCAGATGGTTCAGTTCTGGCCTGTGGCCTGCTGCTGGAGCCCGGAGGA | 120 | | | |
| Qy 121 | gcagctacagctgcctccgtccctctggacaaatggcccaaacactgagcagcat | 180 | | | |
| Db 121 | GCCAGCTACAGCTGCTGCGCGTCCCTCTTGACAAATGGCCCAACACTGAGCAGGCAT | 180 | | | |
| Qy 181 | ctgggtggccctgccaggttgatgccactgctctgcggccactcctcgtcatcttacc | 240 | | | |
| Db 181 | CTGGGTGGCCCTGCCAGGTGATGCCACTGCTCTGCCGGCCACTCCTGCATCTTTACC | 240 | | | |
| Qy 241 | gtctcagggaactccagttgtgctcccttcccagggcgtggtgcatgcggggatggccat | 300 | | | |
| Db 241 | GTCTCAGGGACTTCCAGTTGTGCCCCCTTCCAGAGGCGGTGGCATGGGGGATGGCCAT | 300 | | | |
| Qy 301 | cactgtgcccaacggggttccactgaagtgcagcagggcgagatcgtcttccaaagatca | 360 | | | |
| Db 301 | CAC7GCTGCCCACGGGGCTTCCACTGCAGTGCACAGCGGGCATCTGCTTCCAAAGATCA | 360 | | | |
| Qy 361 | ggtaacaactccgtgggtggccatccagtgccctgatatcagttcgaatgcgggacttc | 420 | | | |
| Db 361 | GGTAACAACCTCCGTGGGTGCCATCCAGTGGCCCTGATAGTCAAGTTCGAATGCCGGACTTC | 420 | | | |
| Qy 421 | tccacgtgctgtgttatggtcgatggctctctgggggtgctgcccacgtccccaggtcc | 480 | | | |
| Db 421 | TCCACGTGCTGTATTATGGTGGATGGGTCTCTGGGGGTGCTGCCCATGCCACAGCTTCC | 480 | | | |
| Qy 481 | tgctgtaaagacaggggtgactgtgtccgcaggtgctctcgcagctgtgttcacac | 540 | | | |
| Db 481 | TGCTGTGAAGACAGGGGTGACTGTGTCGGCACGGTGTCTTCTGGGACTCTGTTACACACC | 540 | | | |
| Qy 541 | cgctgcatcacaccacgggcaccaccccctggcacaagaagctccctgcccagaggact | 600 | | | |
| Db 541 | CGCTGCATCACACCCACGGGCACCCACCCCTGCAAGAAGCTCCCTGCCCGCAGGAGCT | 600 | | | |
| Qy 601 | aacggggcagtggtctgttcagctcg | 627 | | | |
| Db 601 | AACAGGGCAGTGGGCTTTGTCAGCTCG | 627 | | | |

| | |
|------------|--|
| RESULT | 2 |
| AU121637 | |
| LOCUS | |
| DEFINITION | 730 bp mRNA linear EST 19-OCT-2000 MAMMAL Homo sapiens cDNA clone MAMMA1000615 5', mRNA |
| ACCESSION | AU121637 |
| VERSION | AU121637 |
| KEYWORDS | AU121637.1 GI:10936872 |
| SOURCE | EST. |
| ORGANISM | human. |
| | Homo sapiens |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 730)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isoqai, T.

| TITLE | JOURNAL | COMMENT |
|--|---------|---------|
| HRI human cDNA project Unpublished (2000) Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3951 Fax: 81-438-52-3952 | | i |

fax: 01-430-52-3922
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. 730 |

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMAL1000615"
/clone_lib="MAMMAL"
/tissue_type="mammary gland"
/note="Vector: pWE18SFL3"
119 a 240 c 220 g 148 t 3 others
BASE COUNT
ORIGIN

```

| | | | | |
|-----------------------|--------------|---------------------|------------|-------------|
| Query Match | 100.0%; | Score 627; | DB 9; | Length 730; |
| Best Local Similarity | 100.0%; | Pred. No. 8.6e-135; | | |
| Matches 627: | Conservative | 0; | Mismatches | 0; |
| | Indels | 0; | Gaps | 0; |

| | | | |
|----|-----|---|-----|
| Qy | 1 | cgcaaggcagaccattgtggacccttgtagctgggtggccttaacagcaggctggtgact | 60 |
| | | | |
| Db | 29 | CGCAGGCAGACCAATGTGGACCCTGGTAGCTGGGTGGCTTTAAACAGCAGGCTGGTGGCT | 88 |
| | | | |
| Qy | 61 | ggaacgcgggtgccagatggtcagttctgcctctggcctgtgacctgtgcctggaccccgga | 120 |
| | | | |
| Db | 89 | GGAAACGGGTGCCAGATGGTCAGTTCTGCCCTGTGGCTGCTGCTGGACCCCGGAGGA | 148 |
| | | | |
| Qy | 121 | gccagctacagctgctgccgttccctcttgagacaaatggcccacaaactgagcaggcat | 180 |
| | | | |
| Db | 149 | GCCAGCTACAGCTGCTGCCGTCCTCTGTGACAAATGGCCACAACTAGCAGGCGCAT | 208 |
| | | | |
| Qy | 181 | ctgggtggccctgcccaggttgatgcccaatgctctgcgcgccactctgcatctttacc | 240 |
| | | | |
| Db | 209 | CTGGGTGGCCCTGCCAGGTTGATGCCACTGCTCTGCCGGCCACTCTCTGCATCTTTACC | 268 |
| | | | |
| Qy | 241 | gtctcagggaactccagttgtgcccttcccagaggcgtggcattgcggggatggccat | 300 |
| | | | |
| Db | 269 | GTCTCAGGGACTTCCAGTTGCTGCCCTTCGCAGAGGCGCTGGCATGCGGGATGGCCAT | 328 |
| | | | |
| Qy | 301 | cactgctgccacaggggcttccactgagtcagtcagacagggcgatcgtcttccaaagatca | 360 |
| | | | |
| Db | 329 | CACGTCTGCCACAGGGGCTTCCACTGCAGTCGACAGCGGGCGATCCTGCTTCCAAGATCA | 388 |
| | | | |
| Qy | 361 | ggtaacaaactccgtgggtgccatccagtcgacctgatagtcagttcgaaatgcccgacttc | 420 |
| | | | |
| Db | 389 | GGTAACAACCTCGTGGGTGCCATCCAGTGCCTGTATAGTCAGTTCGAATGCCGACATTC | 448 |
| | | | |
| Qy | 421 | tccacgtgctgttatggtcgatggctctctgggggtgctgccccatgccccaggcttcc | 480 |
| | | | |
| Db | 449 | TCCACGTGCTGTGTATATGCTGATGGCTTCCTGGGGGTGCTGCCCATGCCCAGGCTTCC | 508 |
| | | | |
| Qy | 481 | tgctgtgaagacaggggtgcactgctgtccgcagcgggtgctctctgcgacctgggtcacacc | 540 |
| | | | |
| Db | 509 | TGCTGTGAAGACAGGGTGCACTGCTGTGCCGACAGGTGCTTCTGGACCTGGTTCACACC | 568 |
| | | | |
| Qy | 541 | cgctgcatcacacccacagggcacccacccccctggcaagaagactccctgcccagaggact | 600 |
| | | | |

```

Db 569 CGTGCATCACACCCAGCGGACCCACCCCTGCGCAAGAGCTCCCTGCCAGAGGACT 628
Qy 601 aacaggcagtggtgcttgcagctgc 627
Db 629 AACAGGCGAGTGGCTGTGTCAGCTCG 655

RESULT 3
AUI130279
LOCUS AUI130279 NT2RP3 Homo sapiens cDNA clone NT2RP3000541 5', mRNA
DEFINITION sequence.
ACCESSION AUI130279
VERSION AUI130279.1 GI:10990633
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 760)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.
TITLE HRI human cDNA project
JOURNAL HRI human cDNA project
COMMENT Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1. 760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP3000541"
/clone_lib="NT2RP3"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT 123 a 253 c 230 g 151 t
ORIGIN
Query Match 100.0%; Score 627; DB 9; Length 760;
Best Local Similarity 100.0%; Pred. No. 8.7e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cgcaggagaccatgtggaacctggtgagctgggtggtccttaacagaggctggtgct 60
Db 29 CGCAGGAGACCATGTGGACCTGGTGGCTGGTGGCTTAACAGCAGGCTGGTGGCT 88
Qy 61 ggaacgcggtgccagatgctcctgctgctgctgctgctgctgctgctgctgctgct 120
Db 89 GGAACGCGGTGCCAGATGGTCAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 148
Qy 121 gccagctacagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 180
Db 149 GCCAGCTACAGCTGCTGCCGTCCTCTTCTGGCAAAATGGCCACACACTGAGCAGGAT 208
Qy 181 ctgggtgcccctgccagggttgatgccactgctctgcccggcactcctgctgcttacc 240
Db 209 CTGGGTGCCCCCTGCCAGGTGTGATGCCACTGCTCTGCCGCCACTCTCTGATCTTTACC 268
Qy 241 gtctcaggagacttcagttgctcccttccagagaccgtggcagctcgggagggcact 300
Db 269 GTCTCAGGAGACTTCCAGTTGCTGCCCCCTTCCAGAGGCGCGGGCATCGGGGATGGGCAT 328

```

```

Qy 301 cactgtgcccacgggggttccactgcagtcgacagcgggagatcctgcttccaaagatca 360
Db 329 CACTGTGCTCCACGGGGCTTCCACTGCAGTCGACAGCGGCGATCTGTTCCAAAGATCA 388
Qy 361 ggtaaacactccgtgggtgccatccactgcagtcagtcagtcagtcagtcagtcagtc 420
Db 389 GGTAAACACTCCGTGGGTGCCATCCAGTCGCTGATAGTCAGTCAGTCAGTCAGTCAGTC 448
Qy 421 tccactgtgctgtgtatggtgatggtctcctgggggtgctgctgctgctgctgctgctgct 480
Db 449 TCCACGTGCTGTGTATGTTGTCGATGGCTCCTGGGGGTGCTGCCCATGCCAGGCTTCC 508
Qy 481 tgctgtgagacaggggtgcactgctgctgcagcagcagcagcagcagcagcagcagcagc 540
Db 509 TGTGTGAAGACAGGGTGCACACTGCTGTCCGACGGTGGCTTCCGAGCTGGTTCCACACC 568
Qy 541 cgctgcatcacacccacgggcacccacccctggcaagaagctcctgcccagaggact 600
Db 569 CGCTGCATCACACCCAGCGGACCCACCCCTGCGCAAGAGCTCCCTGCCAGAGGACT 628
Qy 601 aacaggcagtggtgcttgcagctgc 627
Db 629 AACAGGCGAGTGGCTGTGTCAGCTCG 655

RESULT 4
BI771292
LOCUS BI771292 782 bp mRNA linear EST 25-SEP-2001
DEFINITION 60304655F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204172 5',
mRNA sequence.
ACCESSION BI771292
VERSION BI771292.1 GI:15762870
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 782)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1511 row: n column: 13
High quality sequence stop: 775.
FEATURES
source
1. 782
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5204172"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/notes="Organ: Pooled lung and spleen; vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 129 a 261 c 232 g 160 t
ORIGIN

```


| | | | |
|-----------------------|---|--|-----------------|
| Db | 569 | CGCTGCATCACACCCACGGGACCCACCCCTGGCAAGAAAGCTCCCTGCCAGGAGCT | 628 |
| Qy | 601 | aacagggcagtgggccttgtccagctcg | 627 |
| Db | 629 | AACAGGCGAGTGGCCTGTTCAGCTCG | 655 |
| RESULT | 10 | | |
| LOCUS | 880 bp | linear | EST 22-MAY-2001 |
| DEFINITION | 602726638F1 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:4866201 | | 5' |
| ACCESSION | BG823258 | | |
| VERSION | BG823258.1 | | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | 1 (bases 1 to 880) | | |
| TITLE | NIH-MGC http://mgc.nci.nih.gov/ . | | |
| JOURNAL | National Institutes of Health, Mammalian Gene Collection (MGC) | | |
| COMMENT | Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1CMI732 row: 1 column: 10 High quality sequence stop: 855. Location/Qualifiers 1. 880 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4866201" /clone_lib="NIH_MGC_15" /tissue_type="adenocarcinoma cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)" | | |
| BASE COUNT | 151 a | 288 c | 261 g |
| ORIGIN | 180 t | | |
| Query Match | 100.0% | Score 627; | DB 10; |
| Best Local Similarity | 100.0% | Identical 0; | Mismatches 0; |
| Matches 627; | Conservative 0; | Indels 0; | Gaps 0; |
| Qy | 1 | cgcgagcgagacatgtgagacctggtagctgggtggccttaacagcaggcgtgtggtc | 60 |
| Db | 12 | CGCAGGCGAGACCATGTGGACCCCTGGTGAAGCTGGGTGGCTTAAACAGCAGGCGTGTGGCT | 71 |
| Qy | 61 | ggaacgcggtgccagatggcagcttcctgtgacctgtgacctgctgacctgacctgacctg | 120 |
| Db | 72 | GGACAGCGGGTGCCAGATGGTTCAGTTCGCCCTGTGGCTCTGCTGCTGGACCCCGAGGA | 131 |
| Qy | 121 | gccagctacagctgtgctgcctccccctcttgacaaaatggcccaacactgagcaggcat | 180 |
| Db | 132 | CCAGCTACAGCTGCTGCGTCCCTTCTGGACAAATGGCCCAACACTGAGCAGGCAT | 191 |
| Qy | 181 | ctgggtggccctgccaggttgatgcccaactgctctgcgcgccaactcgtcatctttacc | 240 |

| | | | |
|----|-----|--|-----|
| Db | 192 | CTGGGTGGCCCTCCAGGTGATGCCCACTGCTCTCCGGCCCACTCCTGCATCTTTACC | 251 |
| Qy | 241 | gtctcaggagacttccagtgtccttcccagagccgtgcatcgcggtatggccat | 300 |
| Db | 252 | GTCTCAGGGACTTCCAGTGTCTGCCCTCTCCAGAGCCGTGGCATCGGGGATGGCCAT | 311 |
| Qy | 301 | cactgctgcaccaggggtcttccactgcagtcagacagggcgatcctctcttcccagaatca | 360 |
| Db | 312 | CACGTGTGCCACAGGGGCTTCCACTGCAGTGTGACAGCGGGCATCTGCTTCCAAAGATCA | 371 |
| Qy | 361 | ggtaacaaactccgtgggtgcataccagtcgcttgatagtcagttcgaatcccgagcttc | 420 |
| Db | 372 | GGTAACAACCTCCGTGGGTGCCATCCACTGCCCTGATAGTCAGTTCGATGCCCGGACTTC | 431 |
| Qy | 421 | tccacgtgctgtgtatggtcgatggctcctgggggtgctgccccatgccccaggcttc | 480 |
| Db | 432 | TCCACGTGCTGTGTATGSGTGATGGCTCTCGGGGTGCTGCCCCATGCCCCAGGCTTC | 491 |
| Qy | 481 | tgctgtgaagacaggggtgcactgctgtccgcacggtgcctctgcgcacctggttcacacc | 540 |
| Db | 492 | TGCTGTGAAGACAGGGTGCACGTGCTGTCCGACAGGTGCCCTTCTGCGACCTGGTTACACCC | 551 |
| Qy | 541 | cgctgcatacaccccacggygcaccccccctggcaagaagctccctgcccagaggact | 600 |
| Db | 552 | CGCTGCATCACCCACGGGCAACCACCCCTGGCAAGAAGCTCCCTGCCTGCCCAGGACT | 611 |
| Qy | 601 | aacagggcagtgcccttgtccagctcg | 627 |
| Db | 612 | AACAGGGCAGTGGCCTTGTGCAGCTCG | 638 |

| | |
|------------|---|
| RESULT | 11 |
| AL552664 | |
| LOCUS | |
| DEFINITION | 920 bp mRNA linear EST 16-FEB-2001 LTI_NFL005_PL2 Homo sapiens cDNA clone CS0DI067YM23 5 prime, mRNA sequence. |

| | | | | | |
|-----------|--|-------------|--|--|--|
| ACCESSION | AL552664 | | | | |
| VERSION | AL552664.1 | GI:12891784 | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 920) | | | | |
| TITLE | Li, W.B., Gruber, C., Jessee, J. and Polayes, D. | | | | |
| JOURNAL | Full-length cDNA libraries and normalization | | | | |
| COMMENT | Unpublished (2001) | | | | |
| | Contact: | Genoscope | | | |

CONTACT : GENOSCOPE
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source

Location/Qualifiers
1. 920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GS0D1067YM23"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMWSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMWSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850. USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

| BASE COUNT | 159 a | 303 c | 272 g | 184 t | 2 others |
|---|-------|-------|-------|-------|----------|
| http://ruirlengthn.inivirogen.com | | | | | |

Query Match 100.0%; Score 627; DB 9; Length 920;

[illegible]

| | |
|------------|--|
| RESULT | 12 |
| AL561424 | |
| LOCUS | 932 bp mRNA linear EST 16-FEB-2001 |
| DEFINITION | AL561424 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL007YF12 5 prime, mRNA sequence. |

Accession AL561424
Version AL561424.1
Keywords EST.
Source human.

JO ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 932)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
source 1. .932

```

1. 1552
Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL007YF12"

```

/clone_lib="LTI_NFL010_BC2"

/sex="male"

/tissue_type="B cells from Burkitt lymphoma"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 160 a 305 c 277 g 189 t 1 others

ORIGIN

```
Query Match 100.0%; Score 627; DB 9; Length 932;
Best Local Similarity 100.0%; Pred. No. 9.3e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cgcaggcagaccatgtgacccctgtgagctgggtggccttaacacagcaggctgtggtc 60
Db 6 CGCAGGACAGACCATGTGGACCCCTGGTGGCTGGTGGCTTAACAGCAGGCTGTGGCT 65
Qy 61 ggaacgggtgcccagatgtgcaattctgcccctgtggtcctgctgctggaacccggagga 120
Db 66 GGAACGGGTGCCAGATGTCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125
Qy 121 gccagctacagctgctgccccttctggaacaaatggcccaacactgagcagcat 180
Db 126 GCCAGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
Qy 181 ctgggtggccctgcccagctgctgctgctgctgctgctgctgctgctgctgctgcttacc 240
Db 186 CTGGGTGCCCCCTGCCAGTTGATGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTTACC 245
Qy 241 gtctcagggaacttccagttgtgctgctgctgctgctgctgctgctgctgctgctgct 300
Db 246 GTCTCAGGGACTTCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
Qy 301 cactgctccacaggggtccactgcaatgcaatgcaatgcaatgcaatgcaatgcaatgca 360
Db 306 CACTGCTCCACAGGGGTTCACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 365
Qy 361 ggttaacaaactcgtgggtgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 420
Db 366 GGTAAACAACTCCGTGGGTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 425
Qy 421 tccacgtgctgttattgtgctgctgctgctgctgctgctgctgctgctgctgcttcc 480
Db 426 TCCACGTGCTGTATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCC 485
Qy 481 tgcctgaagacaggggtgcaatgctgctgctgctgctgctgctgctgctgctgctgct 540
Db 486 TGCTGTGAAGACAGGGGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
Qy 541 cgcctgcatcacacacaggggtgcaatgcaatgcaatgcaatgcaatgcaatgcaatgca 600
Db 546 CGCTGCAATCACACCCAGGGGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
Qy 601 aacagggtgagtgccctgttccagctcgc 627
Db 606 AACAGGGCAGTGGCCTGTGTCAGCTCG 632
```

RESULT 13

AL548072 986 bp mRNA linear EST 16-FEB-2001
LOCUS AL548072 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1034YC17 5
DEFINITION prime, mRNA sequence.
ACCESSION AL548072
VERSION AL548072.1 GI:12882738

KEYWORDS

EST. human.

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 986)

AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..986

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1034YC17"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 168 a 323 c 290 g 203 t 2 others

ORIGIN

```
Query Match 100.0%; Score 627; DB 9; Length 986;
Best Local Similarity 100.0%; Pred. No. 9.5e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cgcaggcagaccatgtgacccctgtgagctgggtggccttaacacagcaggctgtggtc 60
Db 6 CGCAGGACAGACCATGTGGACCCCTGGTGGCTGGTGGCTTAACAGCAGGCTGTGGCT 65
Qy 61 ggaacgggtgcccagatgtgcaattctgcccctgtggtcctgctgctgctgctgctgct 120
Db 66 GGAACGGGTGCCAGATGTCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125
Qy 121 gccagctacagctgctgccccttctggaacaaatggcccaacactgagcagcat 180
Db 126 GCCAGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
Qy 181 ctgggtggccctgcccagctgctgctgctgctgctgctgctgctgctgctgctgcttacc 240
Db 186 CTGGGTGCCCCCTGCCAGTTGATGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTTACC 245
Qy 241 gtctcagggaacttccagttgtgctgctgctgctgctgctgctgctgctgctgctgct 300
Db 246 GTCTCAGGGACTTCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
Qy 301 cactgctccacaggggtccactgcaatgcaatgcaatgcaatgcaatgcaatgcaatgca 360
Db 306 CACTGCTCCACAGGGGTTCACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 365
Qy 361 ggttaacaaactcgtgggtgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 420
Db 366 GGTAAACAACTCCGTGGGTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 425
Qy 421 tccacgtgctgttattgtgctgctgctgctgctgctgctgctgctgctgctgcttcc 480
Db 426 TCCACGTGCTGTATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCC 485
Qy 481 tgcctgaagacaggggtgcaatgctgctgctgctgctgctgctgctgctgctgctgct 540
Db 486 TGCTGTGAAGACAGGGGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
Qy 486 TGCTGTGAAGACAGGGGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
```

```

Qy 541 cgtgcatcacccacgggaccacccccctggcaagaagctccctgcccagaggact 600
Db 546 CGCTGCATCACACCCACGGGACCCACCCCTGGCAAGAGACTCCCTGCCAGAGGACT 605
Qy 601 aacgggagcagtgccctgtccagctcg 627
Db 606 AACAGGGCAGTGGCTTGTCCAGCTCG 632

RESULT 14
BM468251
LOCUS AGENCOURT_6475640 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578048
5', mRNA sequence.
ACCESSION BM468251
VERSION BM468251.1 GI:18517293
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: MGC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2332 row: h column: 17
High quality sequence stop: 670.
FEATURES
source
1..991
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5578048"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 170 a 325 c 295 g 201 t
ORIGIN
Query Match 100.0%; Score 627; DB 10; Length 991;
Best Local Similarity 100.0%; Pred. No. 9.5e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 cgcaggcagaccatgtgacccctggtgagctgggtggccttaacagcagggtggtgct 60
Db 13 CGCAGGCAGACCATGTGGACCCCTGGTGCAGCTGGGTGGCTTAACAGCAGGCTGTGGCT 72
Qy 61 ggaacgcggtgccagatggtcagttctgcccctggtcctgctgcctggaccccgaggga 120
Db 73 GGAACGCGGTGCCAGATGTTGAGTTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 132
Qy 121 gccagctacagctgctccctctctgacaaatggcccaaacactgacagcagcat 180
Db 133 GCCAGCTACAGCTGCTGCCGTGCCCTTCTTGACAAATGGCCCAACACTGAGCAGGCAT 192
Qy 181 ctgggtggccctgccagggttgatgcccaactgctctgcggccactctgcatctttacc 240
Db 193 CTGGGTGCCCTGCCAGGTGATGCCCACTGCTCTGCCGCCACTCTCTGATCTTTACC 252

```

```

Qy 241 gtctcagggaacttcagttgctgccccttcccagagggccgtggcagtcgggagtgccat 300
Db 253 GTCTCAGGAGACTTCCAGTGTGTCGCCCTTCCAGAGCCGTGGCATGGGGATGCCCAT 312
Qy 301 cactgctgccacacggggttccactgacgtgcagcagggcgatcctgttccaaagatca 360
Db 313 CACTGCTGCCACACGGGGTTCCTACTGTCAGTGCAGAGCGGCGATCCTGTTCCAAAGATCA 372
Qy 361 ggtacaactccgtgggtgccatccagtcgacctatagctcagttcgaatgccggaactc 420
Db 373 GGTACAACACTCCGTGGGTGCCATCCAGTGCCTGTAGTTCAGTTCGAATGCCCGACTTC 432
Qy 421 tccagctgctgttatgctgagtcgctccctgggggtgctgcccactgccccagcttcc 480
Db 433 TCCAGCTGCTGTATTGTCGTAGTGCCTTGGGGGTGTCGCCCATGCCCCAGGCTTC 492
Qy 481 tgcgtggaagacaggggtgcaactgctgcgcacgggtgcttctgtgcagactggtcacacc 540
Db 493 TGCTGTGAAGACAGGGTGCCTGCTGTCGCACGCTGCTTCTGCGACTCTGGTTCCACACC 552
Qy 541 cgcgtcatcacacccacggcagccaccccccctggcaagaagctccctgcccagaggact 600
Db 553 CGCTGCATCACACCCACGGGACCCACCCCTGGCAAGAGACTCCCTGCCAGAGGACT 612
Qy 601 aacgggagcagtgccctgtccagctcg 627
Db 613 AACAGGGCAGTGGCTTGTCCAGCTCG 639

RESULT 15
AL542702
LOCUS AL542702 LTI_FL002_PL1 Homo sapiens cDNA clone CSODE011YC02 5 prime
DEFINITION mRNA sequence.
ACCESSION AL542702
VERSION AL542702.1 GI:12875005
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..1010
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODE011YC02"
/clone_lib="LTI_FL002_PL1"
/lab_host="DH10B"
/notes="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 170 a 327 c 300 g 211 t
ORIGIN
Query Match 100.0%; Score 627; DB 9; Length 1010;
Best Local Similarity 100.0%; Pred. No. 9.6e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 18:22:11 ; Search time 111.72 Seconds
(without alignments)
131.919 Million cell updates/sec

Title: US-09-824-647-16_COPY_796_855

Perfect score: 60
Sequence: 1 tccaaggagacgctaccac.....ctgcgcacacagtggcgat 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|------------------|
| 1 | 60 | 100.0 | 2095 | 4 | US-08-991-862-16 |
| 2 | 58.4 | 97.3 | 1779 | 1 | US-07-668-648-3 |
| 3 | 58.4 | 97.3 | 1779 | 2 | US-08-429-998-3 |
| 4 | 58.4 | 97.3 | 1779 | 2 | US-08-431-333-3 |
| 5 | 58.4 | 97.3 | 1779 | 5 | PCT-US91-02321-3 |
| 6 | 34.4 | 57.3 | 539 | 1 | US-07-668-648-7 |
| 7 | 34.4 | 57.3 | 539 | 2 | US-08-429-998-7 |
| 8 | 34.4 | 57.3 | 539 | 2 | US-08-431-333-7 |
| 9 | 34.4 | 57.3 | 539 | 5 | PCT-US91-02321-7 |
| 10 | 31.2 | 52.0 | 1767 | 1 | US-07-668-648-5 |
| 11 | 31.2 | 52.0 | 1767 | 2 | US-08-429-998-5 |
| 12 | 31.2 | 52.0 | 1767 | 2 | US-08-431-333-5 |
| 13 | 31.2 | 52.0 | 1767 | 5 | PCT-US91-02321-5 |
| 14 | 31.2 | 52.0 | 2137 | 4 | US-08-991-862-1 |
| 15 | 26.4 | 44.0 | 1767 | 1 | US-07-668-648-1 |
| 16 | 26.4 | 44.0 | 1767 | 2 | US-08-429-998-1 |
| 17 | 26.4 | 44.0 | 1767 | 2 | US-08-431-333-1 |
| 18 | 26.4 | 44.0 | 1767 | 5 | PCT-US91-02321-1 |
| 19 | 24.4 | 40.7 | 1377 | 4 | US-09-377-557-17 |
| 20 | 23.4 | 39.0 | 4181 | 1 | US-07-670-611-1 |
| 21 | 23.4 | 39.0 | 4181 | 1 | US-08-220-674-1 |
| 22 | 23.4 | 39.0 | 4181 | 1 | US-08-445-186-1 |
| 23 | 23.4 | 39.0 | 4181 | 1 | US-08-446-549-1 |
| 24 | 23.4 | 39.0 | 4181 | 2 | US-08-446-550-1 |
| 25 | 22.2 | 37.0 | 28958 | 1 | US-08-258-261B-6 |
| 26 | 22.2 | 37.0 | 28958 | 1 | US-08-456-837-6 |
| 27 | 22.2 | 37.0 | 28958 | 1 | US-08-457-342-6 |

28 22.2 37.0 28958 1 US-08-457-646A-6 Sequence 6, Appli
29 22.2 37.0 28958 1 US-08-458-076A-6 Sequence 6, Appli
30 22.2 37.0 28958 1 US-08-764-233A-4 Sequence 4, Appli
31 22.2 37.0 28958 1 US-08-457-335A-6 Sequence 6, Appli
32 22.2 37.0 28958 1 US-08-729-214-6 Sequence 6, Appli
33 22.2 37.0 28958 3 US-09-028-934-6 Sequence 6, Appli
34 22.2 37.0 43377 1 US-08-764-233A-1 Sequence 1, Appli
35 22.2 37.0 441529 4 US-09-103-840A-1 Sequence 1, Appli
36 21.8 36.3 252 4 US-08-990-823-73 Sequence 73, Appli
37 21.8 36.3 1881 4 US-09-434-288-5 Sequence 5, Appli
38 21.8 36.3 7653 2 US-08-394-189B-1 Sequence 1, Appli
39 21.8 36.3 7653 3 US-08-258-287B-1 Sequence 1, Appli
40 21.8 36.3 7653 3 US-08-368-704C-1 Sequence 1, Appli
41 21.8 36.3 7653 5 PCT-US93-05701-18 Sequence 18, Appli
42 21.8 36.3 7653 5 PCT-US93-05705-1 Sequence 3, Appli
43 21.8 36.3 15664 1 US-08-402-282-3 Sequence 3, Appli
44 21.8 36.3 15664 1 US-08-508-004-3 Sequence 3, Appli
45 21.8 36.3 15664 1 US-08-402-066-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-991-862-16
; Sequence 16, Application US/08991862
; Patent No. 6309826
; GENERAL INFORMATION:
; APPLICANT: Serrero, Ginette
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/08/991,862
; EARLIER FILING DATE: 1998-08-17
; EARLIER APPLICATION NUMBER: 08/863,862
; EARLIER FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Human GP88 cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(1791)
; OTHER INFORMATION: Nucleotide sequence of human granulin/epithelin
; OTHER INFORMATION: precursor (human GP88). Human Granulin Genebank
; OTHER INFORMATION: M75161.
US-08-991-862-16

Query Match 100.0%; Score 60; DB 4; Length 2095;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tccaaggagacgctaccacgacccctcctcactaagctgctgcgcacacagtggcgat 60
|||||
Db 796 tccaaggagacgctaccacgacccctcctcactaagctgctgcgcacacagtggcgat 855

RESULT 2
US-07-668-648-3
; Sequence 3, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/668,648
FILING DATE: 19910819
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1779 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORGANISM: Homo sapiens
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1779
US-07-668-648-3

Query Match 97.3%; Score 58.4; DB 1; Length 1779;
Best Local Similarity 98.3%; Pred. No. 8.2e-12;
Matches 59; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tccaaggagaacgctaccacgacctctactaagctgctgcgcacacagtgaggcat 60
|||||
Db 784 TCCAAGGAGAGCGCTACCACGCGACCTCTCTACTAGCTGCGCTGCCACACAGTGGGGAT 843

RESULT 3
US-08-429-998-3
Sequence 3, Application US/08429998
Patent No. 5885961
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
TITLE OF INVENTION: MODULATING PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429,998
FILING DATE: 27-APR-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/668,648
FILING DATE: 13-MAR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1779 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORGANISM: Homo sapiens
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1779
US-08-429-998-3

Query Match 97.3%; Score 58.4; DB 2; Length 1779;
Best Local Similarity 98.3%; Pred. No. 8.2e-12;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaaggagaacgctaccacgacctctactaagctgctgcgcacacagtgaggcat 60
|||||
Db 784 TCCAAGGAGAGCGCTACCACGCGACCTCTCTACTAGCTGCGCTGCCACACAGTGGGGAT 843

RESULT 4
US-08-431-333-3
Sequence 3, Application US/08431333
Patent No. 5965723
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
TITLE OF INVENTION: MODULATING PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,333
FILING DATE: 27-APR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/668,648
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1779 base pairs
TYPE: nucleic acid
STRANDEDNESS: single


```

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1779
US-08-431-333-3

Query Match          97.3%; Score 58.4; DB 2; Length 1779;
Best Local Similarity 98.3%; Pred. No. 8.2e-12;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tccaaggagaacgctaccacgagcctcctcactaagctgcctgcgcacacagtgggcgat 60
|||||
Db 784 TCCAAGGAGAGACGCTACCAAGGACCTCTCTACTAAGCTGCTGCCACACACAGTGGGGGAT 843

RESULT 5
PCT-US91-02321-3
; Sequence 3, Application PC/TUS9102321
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US91/02321
; FILING DATE: 19910403
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Foot, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: ON0071A-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)728-4800
; TELEFAX: (206)448-4775
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1779 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1779
PCT-US91-02321-3

Query Match          97.3%; Score 58.4; DB 5; Length 1779;
Best Local Similarity 98.3%; Pred. No. 8.2e-12;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tccaaggagaacgctaccacgagcctcctcactaagctgcctgcgcacacagtgggcgat 60
|||||
Db 784 TCCAAGGAGAGACGCTACCAAGGACCTCTCTACTAAGCTGCTGCCACACACAGTGGGGGAT 843
```

```

|||||
Db 784 TCCAAGGAGAGACGCTACCAAGGACCTCTCTACTAAGCTGCTGCCACACACAGTGGGGGAT 843

RESULT 6
US-07-668-648-7
; Sequence 7, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/668,648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..537
US-07-668-648-7

Query Match          57.3%; Score 34.4; DB 1; Length 539;
Best Local Similarity 85.0%; Pred. No. 0.0015;
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Oy 1 tccaaggagaacgctaccacgagcctcctcactaagctgcctgcgcacacagtgggcgat 60
|||||
Db 103 TCCAAGGAGAGACGCT---ACGGACCTCTCTACCAAGCTGCCCGCCACACACAGTGCAGGAT 159

RESULT 7
US-08-429-998-7
; Sequence 7, Application US/08429998
; Patent No. 5885961
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
```

ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429-998
FILING DATE: 27-APR-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/668,648
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..537
US-08-429-998-7

Query Match 57.3%; Score 34.4; DB 2; Length 539;
Best Local Similarity 85.0%; Pred. No. 0.0015;
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
QY 1 tccaaggaagcgtaccacgacctcactaaagtcgctgcgcacacagtggcgat 60
|||||
DB 103 TCCAGGAGAGCGCT---ACGGACCTCTCCACCAAGCTGCCCGCACACAGTGCAGGAT 159
RESULT 8
US-08-431-333-7
Sequence 7, Application US/08431333
Patent No. 5965723
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
APPLICANT: PLOWMAN, Gregory D.
TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH
TITLE OF INVENTION: MODULATING PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/431.333
FILING DATE: 27-APR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/668,648
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..537
US-08-431-333-7
Query Match 57.3%; Score 34.4; DB 2; Length 539;
Best Local Similarity 85.0%; Pred. No. 0.0015;
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
QY 1 tccaaggaagcgtaccacgacctcactaaagtcgctgcgcacacagtggcgat 60
|||||
DB 103 TCCAGGAGAGCGCT---ACGGACCTCTCCACCAAGCTGCCCGCACACAGTGCAGGAT 159
RESULT 9
PCT-US91-02321-7
Sequence 7, Application PC/TUS9102321
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
APPLICANT: PLOWMAN, Gregory D.
TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH
TITLE OF INVENTION: MODULATING PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Bristol-Myers Squibb Company
STREET: 3005 First Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02321
FILING DATE: 19910403
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Poor, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE/DOCKET NUMBER: ON0071A-PC
TELEPHONE: (206)728-4800
TELEFAX: (206)448-4775
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs

TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..537
PCT-US91-02321-7

Query Match 57.3%; Score 34.4; DB 5; Length 539;
Best Local Similarity 85.0%; Pred. No. 0.0015;
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
Qy 1 tccaaggaacgtaccacgacctctactaagctgctgcgcacacagtgaggcgat 60
Db 103 TCCAGGAGAACGCT---ACGACCTCTCACCACAGCTGCCCGCACACACAGTGCAGGAT 159

RESULT 10
US-07-668-648-5
; Sequence 5, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Flowman, Gregory D.
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/668,648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MISTOCK, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1767 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1767
US-07-668-648-5

Query Match 52.0%; Score 31.2; DB 1; Length 1767;
Best Local Similarity 75.0%; Pred. No. 0.021;

Matches 39; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 8 agaagctaccacgacctctactaagctgctgcgcacacagtgaggcg 59
Db 785 AGAACTACACACGAGTCTCCTGACCAAGCTGCTGGATACCCAGTGAAGGA 836

RESULT 11
US-08-429-998-5
; Sequence 5, Application US/08429998
; Patent No. 5885961
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Flowman, Gregory D.
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,998
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/668,648
; FILING DATE: 13-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MISTOCK, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1767 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1767
US-08-429-998-5

Query Match 52.0%; Score 31.2; DB 2; Length 1767;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 39; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 8 agaagctaccacgacctctactaagctgctgcgcacacagtgaggcg 59
Db 785 AGAACTACACACGAGTCTCCTGACCAAGCTGCTGGATACCCAGTGAAGGA 836

RESULT 12
US-08-431-333-5
; Sequence 5, Application US/08431333
; Patent No. 5965723
; GENERAL INFORMATION:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02321
FILING DATE: 19910403
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Poor, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE CHECKET NUMBER: ON0071A-DC

```

; REFERENCE/DOCKET NUMBER: ON007/IA-P-
;
; TELECOMMUNICATION INFORMATION:
;
;   TELEPHONE: " (206)728-4800
;
;   TELEFAX: (206)448-4775
;
; INFORMATION FOR SEQ ID NO: 5:
;
;   SEQUENCE CHARACTERISTICS:

```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1767 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:

```

ORIGINAL SOURCE: Mus musculus
ORGANISM: Mus musculus
TISSUE TYPE: Kidney
FEATHER:

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1767
```

PCT-US91-02321-5

Query Match 52.0%;
Best Local Similarity 75.0%;
Matches 39; Conservative

Qy 8 agaacgctaccacggacctcctc
||||| ||||||| |||||

Db 785 AGAACTACACCGGATCTCCTG

RESULT 14
US-08-991-862-1
; Sequence 1, Application US/08999

; Patent No. 6309826
 ; GENERAL INFORMATION:
 ; APPLICANT: Serrero, Ginette

```

; ; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; ;
; ; FILE REFERENCE: Z9996.488/P001-A
; ;
; ; CURRENT APPLICATION NUMBER: US/08/991,862
; ;

```

; CURRENT FILING DATE: 1998-08-11
 ;
 ; EARLIER APPLICATION NUMBER: 08
 ;
 ; EARLIER FILING DATE: 1997-05-22

```

; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

```

```

; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Mouse epithelin/gran

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23)..(1789)

```

```

; OTHER INFORMATION: The sequen
; OTHER INFORMATION: mouse gran
; OTHER INFORMATION: instead of

```

; OTHER INFORMATION: (position
US-08-991-862-1

| | |
|-----------------------|-------|
| Query Match | 52.0% |
| Best Local Similarity | 75.0% |

Matches 39; Conservative

Qy 8 agaagctaccacggacctcctc

| | | | | |
|-----------------------|--------------|-------------------|-----------|--------------|
| Query Match | 52.0% | Score 31.2; | DB 4; | Length 2137; |
| Best Local Similarity | 75.0% | Pred. No. 0.021; | | |
| Matches 39: | Conservative | 0: Mismatches 13: | Indels 0: | Gaps |

QY 8 agaacgctaccacaggaccctcctactaagctgcctgcgcacacagtggcgca 59
 matches 39; conservative 0; mismatches 13; indels 0; gaps

Db 807 agaactacacgagatctctgacaaagctgctggatataccacagtgaagga 858
||||| ||||||| ||||| || ||||||| || ||||||| || ||||||| ||

RESULT 15

US-07-668-648-1

; Sequence 1, Application US/07668648

; Patent No. 5416192

; GENERAL INFORMATION:

; APPLICANT: Shoyab, Mohammed

; APPLICANT: Plozman, Gregory D.

; TITLE OF INVENTION: EPIRHELINS: NOVEL CYSTEINE-RICH GROWTH

; TITLE OF INVENTION: MODULATING PROTEINS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/668,648

; FILING DATE: 19910819

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 5624-161-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)790-9090

; TELEFAX: (212) 869-9741

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1767 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1767

US-07-668-648-1

Query Match

44.0%; Score 26.4; DB 1; Length 1767;

Best Local Similarity 69.2%; Pred. No. 0.96;

Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 8 agaacgtaccacgagctctcactaagctgcctgcgcacacagtgggcga 59

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

Db 785 AGGACTACACACAGATCTCATGACCAAGCTGCCTGGATACCCACTGAATGA 836

Search completed: September 5, 2002, 18:22:14

Job time: 13337 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 18:30:01 ; Search time 457.59 seconds
(without alignments)
225.125 Million cell updates/sec

Title: US-09-824-647-16_COPY_796_855

Perfect score: 60

Sequence: 1 tccaaggagaacgtaccac.....ctgcgcacacagtggcgat 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 60 | 100.0 | 2095 | AAV82825 | Human GP88 autocri |
| 2 | 60 | 100.0 | 2124 | AAQ56794 | Granulin coding se |
| 3 | 60 | 100.0 | 2198 | AAQ49052 | Granulin coding se |
| 4 | 58.4 | 97.3 | 757 | AAH07141 | Human cDNA clone (|
| 5 | 58.4 | 97.3 | 1630 | AAH16370 | Human cDNA sequenc |
| 6 | 58.4 | 97.3 | 1779 | AAQ14339 | Human epithelin pr |
| 7 | 58.4 | 97.3 | 2432 | AAQ78180 | Human cancer assoc |
| 8 | 34.4 | 57.3 | 539 | AAQ14952 | Bovine epithelin p |
| 9 | 31.2 | 52.0 | 1767 | AAQ14340 | Mouse epithelin pr |

| | | | | | | |
|----|------|------|------|----|----------|--------------------|
| 10 | 31.2 | 52.0 | 2137 | 20 | AAV82824 | Mouse GP88 autocri |
| 11 | 26.4 | 44.0 | 1767 | 12 | AAQ14338 | Rat epithelin prec |
| 12 | 24.6 | 41.0 | 423 | 22 | ABA44981 | Human breast cell |
| 13 | 24.6 | 41.0 | 423 | 22 | ABA55457 | Human foetal liver |
| 14 | 24.6 | 41.0 | 423 | 22 | ABA25175 | Probe #3641 for ge |
| 15 | 24.6 | 41.0 | 423 | 22 | AAK03696 | Human brain expres |
| 16 | 24.6 | 41.0 | 423 | 22 | AAK29156 | Human bone marrow |
| 17 | 24.6 | 41.0 | 423 | 22 | AAI13744 | Probe #3677 for ge |
| 18 | 24.6 | 41.0 | 423 | 22 | AAI35104 | Probe #3790 used t |
| 19 | 24.6 | 41.0 | 423 | 22 | AAI03625 | Probe #3616 used t |
| 20 | 24.6 | 41.0 | 3512 | 20 | AAV69395 | H. contortus PGP-O |
| 21 | 24.4 | 40.7 | 100 | 17 | AAI30909 | Primer F5 for 80 k |
| 22 | 24.4 | 40.7 | 1377 | 22 | AAI70511 | Rice cryptophan de |
| 23 | 24.4 | 40.7 | 2203 | 17 | AAI30869 | Engineered 80 kd p |
| 24 | 23.8 | 39.7 | 1909 | 21 | AAC49855 | Arabidopsis thalia |
| 25 | 23.4 | 39.0 | 159 | 22 | ABA70391 | Human foetal liver |
| 26 | 23.4 | 39.0 | 159 | 22 | AAK18628 | Human brain expres |
| 27 | 23.4 | 39.0 | 159 | 22 | AAK44555 | Human bone marrow |
| 28 | 23.4 | 39.0 | 159 | 22 | AAI50540 | Probe #19226 used |
| 29 | 23.4 | 39.0 | 441 | 23 | AAS69469 | DNA encoding novel |
| 30 | 23.4 | 39.0 | 447 | 23 | AAS69468 | DNA encoding novel |
| 31 | 23.4 | 39.0 | 488 | 22 | ABA57774 | Human foetal liver |
| 32 | 23.4 | 39.0 | 488 | 22 | AAK05839 | Human brain expres |
| 33 | 23.4 | 39.0 | 488 | 22 | AAK31472 | Human bone marrow |
| 34 | 23.4 | 39.0 | 563 | 22 | AAI37357 | Probe #6043 used t |
| 35 | 23.4 | 39.0 | 563 | 22 | AAS05614 | Mammalian vestibu |
| 36 | 23.4 | 39.0 | 1550 | 21 | AAC76777 | Human cDNA sequenc |
| 37 | 23.4 | 39.0 | 1553 | 22 | AAH13982 | Human ORFX ORF2332 |
| 38 | 23.4 | 39.0 | 1642 | 22 | AAO8067 | Human SPOR TRAF-pr |
| 39 | 23.4 | 39.0 | 1642 | 22 | AAH02922 | Human shear stress |
| 40 | 23.4 | 39.0 | 2012 | 24 | AAS18136 | Human DNAX cytokin |
| 41 | 23.4 | 39.0 | 2141 | 23 | AAS72373 | DNA encoding novel |
| 42 | 23.4 | 39.0 | 2789 | 21 | AAF18162 | Lung cancer associ |
| 43 | 23.4 | 39.0 | 3978 | 23 | ABL14929 | Drosophila melanog |
| 44 | 23.4 | 39.0 | 4181 | 13 | AAQ29345 | MCC gene of chromo |
| 45 | 23.4 | 39.0 | 6557 | 23 | ABL14928 | Drosophila melanog |

ALIGNMENTS

RESULT 1

AAV82825
ID AAV82825 standard; cDNA; 2095 BP.

XX AAV82825;

XX 15-MAR-1999 (first entry)

DE Human GP88 autocrine growth factor cDNA.

XX GP88; granulin; epithelin; human; growth factor; autocrine; tumour;
XX cancer; viral infection; antagonist; therapy; diagnosis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 13..1794

XX FT /*tag= a

XX WO9852607-A1.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-US10555.

XX 16-DEC-1997; 97US-0991862.

XX 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

XX Serrero G;

DR WPI; 1999-045276/04.
 DR P-PSDB; AAW85475.
 XX Composition containing antagonist of growth factor GP88 - useful for
 XX treating cancer and viral diseases and also for diagnosing disease
 PT from altered GP88 expression
 PT
 XX
 XX Disclosure; Fig 9A; 86pp; English.
 XX
 CC This human cDNA sequence includes a coding region for GP88 (see
 CC AAW85474), an 88 kDa glycoprotein autocrine growth factor and
 CC epithelin/granulin precursor that is expressed in a tightly
 CC regulated manner in normal cells, is overexpressed and unregulated
 CC in highly tumorigenic cells derived from normal cells, and which
 CC acts as a stringently required growth stimulator for the
 CC tumorigenic cells. Inhibition of GP88 expression or action in the
 CC tumorigenic cells results in an inhibition of the tumorigenic
 CC properties of the overproducing cells. Antagonists to GP88 are
 CC used to treat diseases associated with increased expression of
 CC GP88, particularly cancer but also viral infections. Fragments of
 CC GP88 are used to raise specific antibodies (used as antagonists,
 CC as diagnostic reagents and for delivering toxins or other
 CC compounds to GP88-expressing cells) and to screen for antibodies.
 CC Antisense oligonucleotides can also be used as antagonists.
 CC Methods are provided for diagnosing disease, or determining
 CC susceptibility to disease, resulting from altered GP88 activity.
 XX
 XX Sequence 2095 BP; 364 A; 681 C; 625 G; 425 T; 0 other;

Query Match 100.0%; Score 60; DB 20; Length 2095;
 Best Local Similarity 100.0%; Pred. No. 2.2e-12;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 796 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 855

RESULT 2
 AAQ56794
 ID AAQ56794 standard; DNA; 2124 BP.
 XX
 AC AAQ56794;
 XX
 XX 22-APR-1994 (first entry)
 XX
 DE Granulin coding sequence.
 XX
 KW Granulin; keratinocytes; wound healing; inhibition; peptide;
 KW granulocytes; leucocytes; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT misc_difference 1379..1381
 FT /tag- a
 FT /transl_except- ATG encodes Valine.
 FT misc_difference 1640..1642
 FT /tag- b
 FT /transl_except- CAG encodes Glycine.
 FT
 XX WO9315195-A.
 PN
 XX
 XX 05-AUG-1993.
 PD
 XX 28-FEB-1992; 92WO-CA00089.
 PF
 XX 03-FEB-1992; 92US-0829233.
 PR
 XX (SOLO/) SOLOMON S.
 PA
 XX Solomon S;
 PI

XX WPI; 1993-320328/40.
 DR P-PSDB; AAR48673.
 XX
 XX New cystine rich granulin peptide(s) from leucocyte(s) - are
 PT keratinocyte inhibitors useful topically for wound healing
 PT
 XX
 XX Disclosure; Figure 4c; 53pp; English.
 XX
 CC The granulin inhibits keratinocytes and is useful in formulations
 CC for promoting the healing of wounds.
 XX
 XX Sequence 2124 BP; 383 A; 685 C; 630 G; 426 T; 0 other;
 SQ
 Query Match 100.0%; Score 60; DB 14; Length 2124;
 Best Local Similarity 100.0%; Pred. No. 2.2e-12;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 809 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 868
 RESULT 3
 AAQ49052
 ID AAQ49052 standard; DNA; 2198 BP.
 XX
 AC AAQ49052;
 XX
 XX 22-APR-1994 (first entry)
 DT
 XX Granulin coding sequence.
 DE
 XX Granulin; keratinocytes; wound healing; inhibition; peptide;
 KW granulocytes; leucocytes; ss.
 KW
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT misc_difference 1453..1455
 FT /tag- a
 FT /transl_except- ATG encodes Valine.
 FT misc_difference 1714..1716
 FT /tag- a
 FT /transl_except- CAG encodes Glycine.
 FT
 XX WO9315195-A.
 PN
 XX
 XX 05-AUG-1993.
 PD
 XX 28-FEB-1992; 92WO-CA00089.
 PF
 XX 03-FEB-1992; 92US-0829233.
 PR
 XX (SOLO/) SOLOMON S.
 PA
 XX Solomon S;
 PI
 XX WPI; 1993-320328/40.
 DR P-PSDB; AAR48673.
 XX
 XX New cystine rich granulin peptide(s) from leucocyte(s) - are
 PT keratinocyte inhibitors useful topically for wound healing
 PT
 XX
 XX Disclosure; Figure 4c; 53pp; English.
 XX
 CC The granulin inhibits keratinocytes and is useful in formulations
 CC for promoting the healing of wounds.
 XX
 XX Sequence 2198 BP; 398 A; 708 C; 646 G; 446 T; 0 other;
 SQ

Query Match 100.0%; Score 60; DB 14; Length 2198;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tccaagagaagcgtaccacgacctccctcactaagctgctgcgcacacagtggcgat 60
|||||
Db 883 tccaagagaagcgtaccacgacctccctcactaagctgctgcgcacacagtggcgat 942

RESULT 4
AAH07141
ID AAH07141 standard; cDNA; 757 BP.
XX
AC AAH07141;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:3976.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 3976; 2537pp + CD ROM; English.
XX

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 757 BP; 139 A; 232 C; 217 G; 165 T; 4 other;

Query Match 97.3%; Score 58.4; DB 22; Length 757;
Best Local Similarity 98.3%; Pred. No. 7.2e-12;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaagagaagcgtaccacgacctccctcactaagctgctgcgcacacagtggcgat 60
|||||
Db 328 tccaagagaagcgtaccacgacctccctcactaagctgctgcgcacacagtggcgat 387

RESULT 5
AAH16370
ID AAH16370 standard; cDNA; 1630 BP.
XX
AC AAH16370;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:15309.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 15309; 2537pp + CD ROM; English.
XX

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX Sequence 1630 BP; 285 A; 517 C; 489 G; 339 T; 0 other;
 SQ Query Match 97.3%; Score 58.4; DB 22; Length 1630;
 Best Local Similarity 98.3%; Pred. No. 8.2e-12;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 60
 |||||
 Db 328 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 387
 |||||

RESULT 6
 ID AAQ14339 standard; DNA; 1779 BP.
 XX
 AC AAQ14339;
 XX
 DT 17-JAN-1992 (first entry)
 DE Human epithelin precursor.
 XX
 KW ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT misc_RNA 41..1819
 FT /*tag= a
 FT /note= "claim 9, page 54"
 XX
 XX WO9115510-A.
 XX
 XX 17-OCT-1991.
 XX
 XX 03-APR-1991; 91WO-US02321.
 XX
 XX 13-MAR-1991; 91US-0083796.
 XX
 XX 03-APR-1990; 90US-0504508.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIB.
 XX
 XX Shoyab M, Plovman GD;
 XX
 XX WPI; 1991-325168/44.
 XX
 XX P-PSDB; AAR14326.
 XX
 XX New cysteine-rich growth modulating proteins, epithelins - useful
 XX as inhibitors of neoplastic cell growth and to promote wound
 XX healing and treat psoriasis
 XX
 XX Disclosure; Fig 22; 97pp; English.
 XX
 XX ET-1 and ET-2 were isolated from rat kidneys and their amino acid
 XX sequences determined. A full length rat ET cDNA (AAQ14338) was obt'd.
 XX by screening a rat kidney cDNA library in lambda gt10 with PCR
 XX generated ET probes. These probes were also used to obtain the mouse
 XX ET gene (AAQ14340) from a mouse T-cell genomic library.
 XX
 XX ET DNA was also obtained from human sources (AAQ14339).
 XX
 XX An anti-sense ribonucleic acid molecule complementary to the
 XX indicated fragment in the features is also claimed.
 XX
 XX ET-1 is a bifunctional growth regulator, capable of stimulating
 XX the growth of some cell types while inhibiting the growth of others.
 XX
 XX ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory
 XX bioactivity. In contrast, however, ET-2 is apparently not capable of
 XX eliciting the growth stimulatory activity characteristic of ET-1 and,
 XX in fact, antagonises this ET-1 activity.
 XX
 XX See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.
 XX
 XX Sequence 1779 BP; 304 A; 578 C; 547 G; 350 T; 0 other;

Query Match 97.3%; Score 58.4; DB 12; Length 1779;
 Best Local Similarity 98.3%; Pred. No. 8.3e-12;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 60
 |||||
 Db 784 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 843
 |||||

RESULT 7
 ID AAC78180 standard; cDNA; 2432 BP.
 XX
 AC AAC78180;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:574.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antisthmatic; antirheumatic; antithratic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 XX Homo sapiens.
 XX
 XX WO200005350-A1.
 XX
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05882.
 XX
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX
 XX WPI; 2000-587533/55.
 XX
 XX P-PSDB; AAB43971.
 XX
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 XX useful for treating or diagnosing e.g. cancer -
 XX
 XX Claim 1; Page 1097-1098; 2352pp; English.
 XX
 XX AAC78180 encode the human cancer associated proteins given
 XX in AAB4398 to AAB44239. The proteins can have activities based on the
 XX tissues and cells the genes are expressed in. Example of activities
 XX include: cytostatic; proliferative; vulnery; immunomodulator;
 XX antidiabetic; antisthmatic; antirheumatic; antithratic;
 XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 XX nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 XX polynucleotides and polypeptides can be used for preventing, treating or
 XX ameliorating medical conditions and diagnosing pathological conditions.
 XX
 XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 XX the present invention may be used to treat immune disorders by activating
 XX or inhibiting the proliferation, differentiation or mobilisation of
 XX immune cells to treat disorders of haematopoietic cells, autoimmune
 XX disorders, allergic reactions, graft versus host disease and organ
 XX rejection, modulate haemostatic or thrombolytic activity, modulate
 XX inflammation, cancers, cardiovascular disorders, neurological disease and
 XX bacterial or viral infections. The peptides, nucleotides, antibodies,
 XX agonists and antagonists may be also be used in drug screens. AAC78449 to
 XX AAC78457 and AAB44240 represent sequences used in the exemplification of
 XX the present invention.

XX SQ Sequence 2432 BP; 460 A; 764 C; 707 G; 496 T; 5 other;

Query Match 97.3%; Score 58.4; DB 21; Length 2432;
Best Local Similarity 98.3%; Pred. No. 8.7e-12;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaagagaaagcgtaccacgacgtccctcactaagctgcctgcgcacacagtgggcgat 60
|||||
Db 971 tccaagagaaagcgtaccacgacgtccctcactaagctgcctgcgcacacagtgggcgat 1030

RESULT 8

AAQ14952

ID AAQ14952 standard; DNA; 539 BP.

XX AC AAQ14952;

XX DT 17-JAN-1992 (first entry)

XX DE Bovine epithelin precursor (partial).

XX KW ET; growth regulation; inhibition; stimulation; ss.

XX OS Bos taurus.

XX PN WO9115510-A.

XX PD 17-OCT-1991.

XX PF 03-APR-1991; 91WO-US02321.

XX PR 13-MAR-1991; 91US-0083796.

XX PR 03-APR-1990; 90US-0504508.

XX PA (BRIM) BRISTOL-MYERS SQUIB.

XX PI Shoyab M, Plowman GD;

XX WPI; 1991-325168/44.

XX DR P-PSDB; AARI5426.

XX New cysteine-rich growth modulating proteins, epithelins - useful

XX PT as inhibitors of neoplastic cell growth and to promote wound

XX PT healing and treat psoriasis

XX PS Disclosure; Fig 24; 97pp; English.

XX The epithelins appear to comprise several distinct members sharing

XX CC significant structural homology. Two members of the epithelin family,

XX CC EP-1 and EP-2, have been purified from natural sources, and cDNAs

XX CC encoding these and several other members of the epithelin family have

XX CC been isolated from rat (AAQ14338), human (AAQ14339), bovine, murine

XX CC (AAQ14340) and chicken (AAQ14953).

XX CC ET-1 is a bifunctional growth regulator, capable of stimulating

XX CC the growth of some cell types while inhibiting the growth of others.

XX CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory

XX CC bioactivity. In contrast, however, ET-2 is apparently not capable of

XX CC eliciting the growth stimulatory activity characteristic of ET-1 and,

XX CC in fact, antagonises this ET-1 activity.

XX CC See also AAQ14338-40, AAQ14952-53, AARI14328-9 and AARI15315-20.

XX SQ Sequence 539 BP; 99 A; 171 C; 163 G; 106 T; 0 other;

Query Match 57.3%; Score 34.4; DB 12; Length 539;

Best Local Similarity 85.0%; Pred. No. 0.0036;

Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 tccaagagaaagcgtaccacgacgtccctcactaagctgcctgcgcacacagtgggcgat 60
|||||
Db 103 tccaagagaaagcgt---acggacctcctcaccgaagctgccgcgcacacagtgcgagat 159

RESULT 9

AAQ14340

ID AAQ14340 standard; DNA; 1767 BP.

XX AC AAQ14340;

XX DT 17-JAN-1992 (first entry)

XX DE Mouse epithelin precursor.

XX KW ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.

XX OS Mus musculus;

XX FH Key Location/Qualifiers

XX FT misc_RNA 8..1774

XX FT /*tag- a

XX FT /note= *claim 29, page 56"

XX PN WO9115510-A.

XX PD 17-OCT-1991.

XX PF 03-APR-1991; 91WO-US02321.

XX PR 13-MAR-1991; 91US-0083796.

XX PR 03-APR-1990; 90US-0504508.

XX PA (BRIM) BRISTOL-MYERS SQUIB.

XX PI Shoyab M, Plowman GD;

XX WPI; 1991-325168/44.

XX DR P-PSDB; AARI14327.

XX New cysteine-rich growth modulating proteins, epithelins - useful

XX PT as inhibitors of neoplastic cell growth and to promote wound

XX PT healing and treat psoriasis

XX PS Disclosure; Fig 23; 97pp; English.

XX ET-1 and ET-2 were isolated from rat kidneys and their amino acid

XX CC sequences determined. A full length rat ET cDNA (AAQ14338) was obtd.

XX CC by screening a rat kidney cDNA library in lambda gt10 with PCR

XX CC generated ET probes. These probes were also used to obtain the mouse

XX CC ET gene (AAQ14340) from a mouse T-cell genomic library.

XX CC ET DNA was also obtained from human sources (AAQ14339).

XX CC An anti-sense ribonucleic acid molecule complementary to the

XX CC indicated fragment in the features is also claimed.

XX CC ET-1 is a bifunctional growth regulator, capable of stimulating

XX CC the growth of some cell types while inhibiting the growth of others.

XX CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory

XX CC bioactivity. In contrast, however, ET-2 is apparently not capable of

XX CC eliciting the growth stimulatory activity characteristic of ET-1 and,

XX CC in fact, antagonises this ET-1 activity.

XX CC See also AAQ14338-40, AAQ14952-53, AARI14328-9 and AARI15315-20.

XX SQ Sequence 1767 BP; 360 A; 503 C; 492 G; 412 T; 0 other;

Query Match 52.0%; Score 31.2; DB 12; Length 1767;

Best Local Similarity 75.0%; Pred. No. 0.064;

Matches 39; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 8 agaactaccacgcgcctcactaagctgcctgcgcacacagtgggcgga 59
|||||
Db 785 agaactaccacgcgcctcactaagctgcctgcgcacacagtggaagga 836

RESULT 10

AAV82824

[illegible]

RESULT 13
ABA55457

```
OS Homo sapiens.
XX WO200157274-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00666.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX Claim 1; SEQ ID No 3641; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 423 BP; 110 A; 126 C; 100 G; 87 T; 0 other;

Query Match 41.0%; Score 24.6; DB 22; Length 423;
Best Local Similarity 70.2%; Pred. No. 13;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 5 aggagaacgtaccacgacctctcactaagctgctgcgcacaca 51
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 aggggaagataccaccagggttaacctcactaattgccctgcgcataca 222

RESULT 15
AAK03696
ID AAK03696 standard; DNA; 423 BP.
AC AAK03696;
XX 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe SEQ ID NO: 3687.
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX Homo sapiens.
XX WO200157275-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00667.
```

```
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX Example 4; SEQ ID NO: 3687; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX SQ Sequence 423 BP; 110 A; 126 C; 100 G; 87 T; 0 other;

Query Match 41.0%; Score 24.6; DB 22; Length 423;
Best Local Similarity 70.2%; Pred. No. 13;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 5 aggagaacgtaccacgacctctcactaagctgctgcgcacaca 51
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 aggggaagataccaccagggttaacctcactaattgccctgcgcataca 222

Search completed: September 5, 2002, 18:30:02
Job time: 10534 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 15:34:05 ; Search time 3259.55 Seconds
(without alignments)
248.445 Million cell updates/sec

Title: us-09-824-647-16_COPY_796_855

Perfect score: 60
Sequence: 1 tcacaggagaacgtaccac.....ctgcgcacacagtgggcgat 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pln: *
16: em_gss_vrt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|---------------------|
| c 1 | 58.4 | 97.3 | 134 | 9 AW062972 | AW062972 IL1-ST004 |
| c 2 | 58.4 | 97.3 | 369 | 10 BF902443 | BF902443 RC4-WT023 |
| c 3 | 58.4 | 97.3 | 393 | 9 AA122084 | AA122084 zm23e01_r |
| c 4 | 58.4 | 97.3 | 394 | 10 BF925544 | BF925544 CM2-WT016 |
| c 5 | 58.4 | 97.3 | 394 | 10 BG004140 | BG004140 CM2-GN016 |
| c 6 | 58.4 | 97.3 | 440 | 9 AI751894 | AI751894 cn12e02.x |
| c 7 | 58.4 | 97.3 | 457 | 10 BF998533 | BF998533 CM2-GN016 |
| c 8 | 58.4 | 97.3 | 460 | 10 BG015966 | BG015966 RC4-GN032 |
| c 9 | 58.4 | 97.3 | 482 | 10 BF998626 | BF998626 CM2-GN016 |
| c 10 | 58.4 | 97.3 | 500 | 10 BG014975 | BG014975 RC4-GN032 |
| c 11 | 58.4 | 97.3 | 553 | 10 BE874824 | BE874824 601488765 |
| c 12 | 58.4 | 97.3 | 557 | 10 BF970308 | BF970308 602273694 |
| c 13 | 58.4 | 97.3 | 570 | 9 AW579074 | AW579074 RC0-CT038 |
| c 14 | 58.4 | 97.3 | 576 | 10 BI196113 | BI196113 602754572 |
| c 15 | 58.4 | 97.3 | 594 | 10 BI159891 | BI159891 602863719 |
| c 16 | 58.4 | 97.3 | 595 | 10 BE253430 | BE253430 601111692 |
| c 17 | 58.4 | 97.3 | 608 | 10 BE302516 | BE302516 ba666003.y |

```

18 58.4 97.3 630 10 BE741541
19 58.4 97.3 632 10 BI835123
20 58.4 97.3 641 10 BG423187
21 58.4 97.3 648 10 BM043523
22 58.4 97.3 657 10 BE742649
23 58.4 97.3 663 9 AU135509
24 58.4 97.3 664 10 BE743017
25 58.4 97.3 665 10 BE336895
26 58.4 97.3 677 10 BG748377
27 58.4 97.3 694 10 BE407219
28 58.4 97.3 696 10 BE258826
29 58.4 97.3 708 10 BI198717
30 58.4 97.3 716 10 BM048423
31 58.4 97.3 719 10 BI193612
32 58.4 97.3 737 10 BM043572
33 58.4 97.3 744 10 BF968260
34 58.4 97.3 757 9 AU134038
35 58.4 97.3 771 10 BI252407
36 58.4 97.3 776 10 BE386839
37 58.4 97.3 802 10 BE407885
38 58.4 97.3 803 10 BI870792
39 58.4 97.3 808 10 BG282501
40 58.4 97.3 813 10 BG753796
41 58.4 97.3 825 10 BF345137
42 58.4 97.3 838 10 BG284875
43 58.4 97.3 840 10 BM045888
44 58.4 97.3 844 10 BE383965
45 58.4 97.3 865 9 AL551474

```

ALIGNMENTS

```

RESULT 1
LOCUS AW062972/c
DEFINITION IL1-ST0041-020899-001-F07 ST0041 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW062972
VERSION AW062972.1 GI:6014357
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 134)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-IL1-ST0041-020
899-001-F07&t3=1999-08-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 134.
Location/Qualifiers
1. 134
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ST0041"
/dev_stage="Adult"
/note="*Organ: stomach; Vector: pUC18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)

```

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 27 a 39 c 34 g 34 t
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 134;
Best Local Similarity 98.3%; Pred. No. 3.8e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60
|||||
Db 100 TCCAAGGAGACGCTACCACGGACCTCTCTCACTAAGCTGCTCGGCACACAGTGGGGGAT 41

RESULT 2
BF902443/c
LOCUS BF902443 369 bp mRNA linear EST 18-JAN-2001
DEFINITION RC4-MT0235-131200-021-b04 MT0235 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF902443
VERSION BF902443.1 GI:12293902
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 369)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4&t2=RC4-MT0235-
131200-021-b04&t3=2000-12-13&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 52
High quality sequence stop: 369.

Location/Qualifiers
1..369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0235"
/dev_stage="Adult"

/note="organ: marrow; Vector: puc18; Site_1: Smar; Site_2:
Smar; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 70 a 113 c 108 g 78 t
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 369;
Best Local Similarity 98.3%; Pred. No. 5.2e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60
|||||
Db 176 TCCAAGGAGACGCTACCACGGACCTCTCTCACTAAGCTGCTCGGCACACAGTGGGGGAT 235

RESULT 4
AA122084
LOCUS AA122084 393 bp mRNA linear EST 23-DEC-1997
DEFINITION zm23e01.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone
IMAGE:526488 5' similar to gb:X62320 GRANULINS PRECURSOR (HUMAN);,
mRNA sequence.
ACCESSION AA122084
VERSION AA122084.1 GI:1678148
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 393)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Travaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
9704478
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

WARNING: There is evidence that suggests that the 384-well parent
plate of this clone contains both human and mouse derived clones.
Thus, the origin of this clone is uncertain. This caution should be
kept in mind should you use this clone.

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1145 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 322.

Location/Qualifiers
1..393
/organism="Homo sapiens"
/db_xref="GDB:3917917"
/db_xref="taxon:9606"
/clone_lib="IMAGE:526488"
/lab_host="Stratagene pancreas (#937208)"
/note="organ: pancreas; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dr. Pancreatic adenocarcinoma cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor sequence: 5'
CTCGAGTATTTTATTTT 3'."

BASE COUNT 68 a 118 c 117 g 89 t
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 393;
Best Local Similarity 98.3%; Pred. No. 5.3e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60
|||||
Db 176 TCCAAGGAGACGCTACCACGGACCTCTCTCACTAAGCTGCTCGGCACACAGTGGGGGAT 235


```

BF925544      394 bp      mRNA      linear      EST 19-JAN-2001
LOCUS          CM2-NT0169-291100-525-c04 NT0169 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION      BF925544
VERSION
KEYWORDS
SOURCE
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS        1 (bases 1 to 394)
               Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
               Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
               Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
               Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
               ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
               Simpson,A.J.
               Shotgun sequencing of the human transcriptome with ORF expressed
               sequence tags
               Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
COMMENT        Contact: Simpson A.J.G.
               Laboratory of Cancer Genetics
               Ludwig Institute for Cancer Research
               Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
               Brazil
               Tel: +55-11-2704922
               Fax: +55-11-2707001
               Email: asimpson@ludwig.org.br
               This sequence was derived from the FAPESP/LICR Human Cancer Genome
               Project. This entry can be seen in the following URL
               (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-NT0169-
               291100-525-c04&t3=2000-11-29&t4=1)
               Seq primer: puc 18 forward
               High quality sequence start: 17
               High quality sequence stop: 394.
               Location/Qualifiers
               1..394
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="NT0169"
               /dev_stage="Adult"
               /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
               Site_2: SmaI; A mini-library was made by cloning products
               derived from ORSTES PCR (U.S. Letters Patent application
               No. 196,716 - Ludwig Institute for Cancer Research)
               profiles into the pUC 18 vector. Reverse transcription of
               tissue mRNA and cDNA amplification were performed under
               low stringency conditions."
               71 a 123 c 118 g 82 t

BASE COUNT    71 a 123 c 118 g 82 t
ORIGIN

Query Match      97.3%; Score 58.4; DB 10; Length 394;
Best Local Similarity 98.3%; Pred. No. 5.3e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60
|||||
Db 205 TCCAAGGAGACGCTACCACGACGCTCTCTCACTAAGCTGCCTGCACACAGTGGGGAT 264

RESULT 5
BG004140
LOCUS          CM2-GN0166-201100-525-c04 GN0166 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION      BG004140
VERSION
KEYWORDS
SOURCE
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS        1 (bases 1 to 394)
               Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
               Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
               Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
               Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
               ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
               Simpson,A.J.
               Shotgun sequencing of the human transcriptome with ORF expressed
               sequence tags
               Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
COMMENT        Contact: Simpson A.J.G.
               Laboratory of Cancer Genetics
               Ludwig Institute for Cancer Research
               Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
               Brazil
               Tel: +55-11-2704922
               Fax: +55-11-2707001
               Email: asimpson@ludwig.org.br
               This sequence was derived from the FAPESP/LICR Human Cancer Genome
               Project. This entry can be seen in the following URL
               (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-NT0169-
               291100-525-c04&t3=2000-11-29&t4=1)
               Seq primer: puc 18 forward
               High quality sequence start: 17
               High quality sequence stop: 394.
               Location/Qualifiers
               1..394
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="NT0169"
               /dev_stage="Adult"
               /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
               Site_2: SmaI; A mini-library was made by cloning products
               derived from ORSTES PCR (U.S. Letters Patent application
               No. 196,716 - Ludwig Institute for Cancer Research)
               profiles into the pUC 18 vector. Reverse transcription of
               tissue mRNA and cDNA amplification were performed under
               low stringency conditions."
               71 a 123 c 118 g 82 t

BASE COUNT    71 a 123 c 118 g 82 t
ORIGIN

Query Match      97.3%; Score 58.4; DB 10; Length 394;
Best Local Similarity 98.3%; Pred. No. 5.3e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60
|||||
Db 205 TCCAAGGAGACGCTACCACGACGCTCTCTCACTAAGCTGCCTGCACACAGTGGGGAT 264

RESULT 6
A1751894
LOCUS          cn12e02.xl Normal Human Trabecular Bone Cells Homo sapiens cDNA
DEFINITION      clone NHTBC_cn12e02 random, mRNA sequence.
ACCESSION      A1751894
VERSION
KEYWORDS
SOURCE
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS        1 (bases 1 to 440)
               Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
               Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey
               ,P.G., Hotchkiss,R.N. and Francomano,C.A.
               SGAP: The Skeletal Genome Anatomy Project
               Unpublished (1997)
               Contact: Libin Jia
               Medical Genetics Branch
               National Human Genome Research Institute

```

```

REFERENCE
AUTHORS        1 (bases 1 to 394)
               Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
               Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
               Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
               Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
               ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
               Simpson,A.J.
               Shotgun sequencing of the human transcriptome with ORF expressed
               sequence tags
               Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
COMMENT        Contact: Simpson A.J.G.
               Laboratory of Cancer Genetics
               Ludwig Institute for Cancer Research
               Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
               Brazil
               Tel: +55-11-2704922
               Fax: +55-11-2707001
               Email: asimpson@ludwig.org.br
               This sequence was derived from the FAPESP/LICR Human Cancer Genome
               Project. This entry can be seen in the following URL
               (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-GN0166-
               201100-525-c04&t3=2000-11-20&t4=1)
               Seq primer: puc 18 forward
               High quality sequence start: 17
               High quality sequence stop: 394.
               Location/Qualifiers
               1..394
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="GN0166"
               /dev_stage="Adult"
               /note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
               ; Site_2: SmaI; A mini-library was made by cloning
               products derived from ORSTES PCR (U.S. Letters Patent
               application No. 196,716 - Ludwig Institute for Cancer
               Research) profiles into the pUC 18 vector. Reverse
               transcription of tissue mRNA and cDNA amplification were
               performed under low stringency conditions."
               71 a 123 c 118 g 82 t

BASE COUNT    71 a 123 c 118 g 82 t
ORIGIN

Query Match      97.3%; Score 58.4; DB 10; Length 394;
Best Local Similarity 98.3%; Pred. No. 5.3e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60
|||||
Db 205 TCCAAGGAGACGCTACCACGACGCTCTCTCACTAAGCTGCCTGCACACAGTGGGGAT 264

RESULT 6
A1751894
LOCUS          cn12e02.xl Normal Human Trabecular Bone Cells Homo sapiens cDNA
DEFINITION      clone NHTBC_cn12e02 random, mRNA sequence.
ACCESSION      A1751894
VERSION
KEYWORDS
SOURCE
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS        1 (bases 1 to 440)
               Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
               Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey
               ,P.G., Hotchkiss,R.N. and Francomano,C.A.
               SGAP: The Skeletal Genome Anatomy Project
               Unpublished (1997)
               Contact: Libin Jia
               Medical Genetics Branch
               National Human Genome Research Institute

```

10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157

Email: libhelix.nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).

Plate: 12 row: e column: 02
Seq primer: -21M13 forward primer (ABI).

FEATURES

source

1. 440
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NHTBC.cml2e02"
/clone_lib="Normal Human Trabecular Bone Cells"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI"
BASE COUNT 81 a 149 c 125 g 85 t
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 440;
Best Local Similarity 98.3%; Pred. No. 5.5e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaagagaacgtaccacgacctctactaagctgctgcgcacacagtgaggat 60
|||||
Db 298 TCCAAGAGAGACGTACCACGACCTCTCTACTAAGCTGCTGCACACAGTGGGGAT 357

RESULT 7

BF998533/c

LOCUS BF998533 457 bp mRNA linear EST 24-JAN-2001
DEFINITION CM2-GN0166-101100-496-f05 GN0166 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF998533
VERSION BF998533.1 GI:12434060
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 457)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2at2-CM2-GN0166-
101100-496-f05at3-2000-11-10at4-1)

Seq primer: puc 18 forward

High quality sequence start: 24

High quality sequence stop: 328.

FEATURES

source

1. 457
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone_lib="GN0166"

/dev_stage="Adult"

/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

86 t

BASE COUNT

ORIGIN

97 a

139 c

135 g

Query Match 97.3%; Score 58.4; DB 10; Length 457;

Best Local Similarity 98.3%; Pred. No. 5.6e-09;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaagagaacgtaccacgacctctactaagctgctgcgcacacagtgaggat 60

|||||

Db 281 TCCAAGAGAGACGTACCACGACCTCTCTACTAAGCTGCTGCACACAGTGGGGAT 222

RESULT 8

BG015966

LOCUS

DEFINITION RC4-GN0323-281200-032-g01 GN0323 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BG015966

VERSION

BG015966.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 460)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,

M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4st2-RC4-GN0323-
281200-032-g01st3-2000-12-28at4-1)

Seq primer: puc 18 forward

High quality sequence start: 103

High quality sequence stop: 460.

Location/Qualifiers

1. 460

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="GN0323"

/dev_stage="Adult"

/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI

; Site_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

Application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the pUC 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

85 a 135 c 149 g 91 t

BASE COUNT

ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 460;
 Best Local Similarity 98.3%; Pred. No. 5.7e-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaaggagaagctaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60
 |||||
 Db 181 TCCAAGGAGAAGCTACCACGCGCTCTCCTACTAGCTGCTGGCGCACACAGTGGGGGAT 240

RESULT 9

BF998626 482 bp mRNA linear EST 24-JAN-2001
 LOCUS CM2-GN0166-101100-496-g05 GN0166 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF998626
 ACCESSION BF998626
 VERSION BF998626.1 GI:12434248
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 482)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-GN0166-
 101100-496-g05&t3=2000-11-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 20

High quality sequence stop: 130.

FEATURES

source

Location/Qualifiers

1..482

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="GN0166"

/dev_stage="Adult"

/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI

; Site_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the puc 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

94 a 142 c 146 g 100 t

BASE COUNT

ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 482;
 Best Local Similarity 98.3%; Pred. No. 5.7e-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaaggagaagctaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60
 |||||
 Db 221 TCCAAGGAGAAGCTACCACGCGCTCTCCTACTAGCTGCTGGCGCACACAGTGGGGGAT 280

RESULT 10

BG014975

LOCUS

RC4-GN0323-211200-021-f01

GN0323 Homo sapiens cDNA, mRNA sequence.

ACCESSION BG014975

VERSION BG014975.1

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 500)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC4&t2=RC4-GN0323-
 211200-021-f01&t3=2000-12-21&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 35

High quality sequence stop: 500.

Location/Qualifiers

1..500

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="GN0323"

/dev_stage="Adult"

/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI

; Site_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the puc 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

103 a 160 c 139 g 98 t

BASE COUNT

ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 500;
 Best Local Similarity 98.3%; Pred. No. 5.7e-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaaggagaagctaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60
 |||||
 Db 140 TCCAAGGAGAAGCTACCACGCGCTCTCCTACTAGCTGCTGGCGCACACAGTGGGGGAT 199

RESULT 11

BE874824

LOCUS

601488765F1 NIH_MGC_69

Homo sapiens cDNA clone IMAGE:3890969 5',

DEFINITION BE874824

mrna sequence.

ACCESSION BE874824

VERSION BE874824.1

KEYWORDS EST.

SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 553)
AUTHORS
TITLE
JOURNAL
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DTG/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9675 row: a column: 18
High quality sequence stop: 540.

FEATURES
source

1. .553
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3890969"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."

BASE COUNT 102 a 178 c 162 g 111 t
ORIGIN

Query Match
Best Local Similarity

97.3%; Score 58.4; DB 10; Length 553;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1

tccaagagacgctaccacgacctctcactaagctgctgcgcacagtgaggat 60
|||||
Db 173 TCCAAGGAGACGCTACCACGGACCTCTCACTAAGCTGCTGGCACACAGTGGGGAT 232

RESULT 12
BF970308

LOCUS 602273694F1 NIH_MGC_84 557 bp mRNA linear EST 22-JAN-2001
DEFINITION
mRNA sequence.
ACCESSION BF970308
VERSION BF970308.1 GI:12337596
KEYWORDS
EST.

SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 557)
AUTHORS
TITLE
JOURNAL
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1004 row: 1 column: 23
High quality sequence stop: 556.

FEATURES
source

1. .557
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4361638"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: adrenal gland; Vector: pCMV-Sport6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally. Oligo-dT
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 105 a 182 c 155 g 115 t
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 557;
Best Local Similarity 98.3%; Pred. No. 5.9e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaagagacgctaccacgacctctcactaagctgctgcgcacagtgaggat 60
|||||
Db 458 TCCAAGGAGACGCTACCACGGACCTCTCACTAAGCTGCTGGCACACAGTGGGGAT 517

RESULT 13
AW579074

LOCUS AW579074 570 bp mRNA linear EST 16-MAR-2000
DEFINITION
RC0-CT0380-070100-031-g04 CT0380 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW579074
VERSION AW579074.1 GI:7254123
KEYWORDS
EST.

SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 570)
HCGP http://www.ludwig.org.br/ORESTES.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0et2-RC0-CT0380-
070100-031-g04&t3=2000-01-07&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 570.
High quality sequence stop: 570.

FEATURES
source

1. .570
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0380"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, '76 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 109 a 183 c 159 g 119 t
ORIGIN

Query Match
Best Local Similarity

97.3%; Score 58.4; DB 9; Length 570;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagcgtaccacggacctctcactaaagctgctgcgcacacagtgggcgat 60
|||||
Db 468 TCCAAGGAGAACGCTACCACGGACCTCCTCACTAAGCTGCTGCGCACACAGTGGGGGAT 527

RESULT 14
BI196113
LOCUS 602754572F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4890472 5',
DEFINITION mRNA sequence.
ACCESSION BI196113
VERSION BI196113.1 GI:14651133
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: L1CM1767 row: 0 column: 17
High quality sequence stop: 575.

FEATURES
Location/Qualifiers
1..576
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4890472"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the adaptor: GGCACGAG(C). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 107 a 189 c 164 g 116 t
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 576;
Best Local Similarity 98.3%; Pred. No. 6e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagcgtaccacggacctctcactaaagctgctgcgcacacagtgggcgat 60
|||||
Db 487 TCCAAGGAGAACGCTACCACGGACCTCCTCACTAAGCTGCTGCGCACACAGTGGGGGAT 546

RESULT 15
BI159891
LOCUS 602863719F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5017733 5',
DEFINITION mRNA sequence.
ACCESSION BI159891
VERSION BI159891.1 GI:14619892
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 594)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: L1CM1830 row: f column: 06
High quality sequence start: 6
High quality sequence stop: 582.

FEATURES
Location/Qualifiers
1..594
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5017733"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the adaptor: GGCACGAG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 119 a 182 c 181 g 112 t
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 594;
Best Local Similarity 98.3%; Pred. No. 6e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagcgtaccacggacctctcactaaagctgctgcgcacacagtgggcgat 60
|||||
Db 95 TCCAAGGAGAACGCTACCACGGACCTCCTCACTAAGCTGCTGCGCACACAGTGGGGGAT 154

Search completed: September 5, 2002, 15:34:08
Job time: 8746 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 15:34:08 ; Search time 3259.55 Seconds

(without alignments)

4438.875 Million cell updates/sec

Title: US-09-824-647-16_COPY_1024_2095

Perfect score: 1072

Sequence: 1 gggcccccaccaggcgccctg.....ataaagttgtcactttctt 1072

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| C 1 | 918 | 85.6 | 1018 | 9 | AL572883 |
| C 2 | 904 | 84.3 | 996 | 9 | AL582634 |
| C 3 | 860.8 | 80.3 | 942 | 9 | AL542659 |
| C 4 | 817.8 | 76.3 | 900 | 9 | AL574944 |
| C 5 | 765 | 71.4 | 887 | 9 | AL576991 |
| C 6 | 760.4 | 70.9 | 810 | 10 | BG684182 |
| C 7 | 757 | 70.6 | 908 | 9 | AL562633 |
| C 8 | 747 | 69.7 | 808 | 9 | AL579730 |
| C 9 | 734.8 | 68.5 | 835 | 9 | AL575030 |
| C 10 | 713 | 66.5 | 780 | 10 | BG763617 |
| C 11 | 695 | 64.8 | 745 | 10 | BG281390 |
| C 12 | 693.4 | 64.7 | 807 | 10 | BE742164 |
| C 13 | 681.2 | 63.5 | 894 | 10 | BE742164 |
| C 14 | 673.2 | 62.8 | 751 | 10 | BG830299 |
| C 15 | 667.6 | 62.3 | 888 | 10 | BG23057 |
| C 16 | 664.6 | 62.0 | 877 | 10 | BG284114 |
| C 17 | 662.2 | 61.8 | 806 | 10 | BG742228 |

| | | | | | |
|------|-------|------|------|----|----------|
| 18 | 658.4 | 61.4 | 745 | 10 | BI562752 |
| 19 | 649 | 60.5 | 706 | 10 | BE743031 |
| 20 | 647.4 | 60.4 | 878 | 10 | BG765435 |
| 21 | 639.8 | 59.7 | 732 | 10 | BE747981 |
| 22 | 639.2 | 59.6 | 1082 | 10 | BM466432 |
| C 23 | 636.8 | 59.4 | 753 | 9 | AL570288 |
| C 24 | 611.4 | 57.0 | 714 | 10 | BM007936 |
| C 25 | 604.6 | 55.4 | 753 | 9 | AW131849 |
| C 26 | 596.8 | 55.7 | 687 | 10 | BI831698 |
| C 27 | 596.6 | 55.7 | 734 | 9 | AI985672 |
| C 28 | 593.6 | 55.4 | 967 | 10 | BG110722 |
| C 29 | 593.4 | 55.4 | 766 | 10 | BI200080 |
| C 30 | 584.8 | 54.6 | 769 | 10 | BI252854 |
| C 31 | 584.6 | 54.5 | 628 | 9 | BE042869 |
| C 32 | 581.8 | 54.3 | 920 | 10 | BG770670 |
| C 33 | 579.6 | 54.1 | 710 | 10 | BG251638 |
| C 34 | 575.6 | 53.7 | 702 | 10 | BF981909 |
| C 35 | 575.2 | 53.7 | 639 | 9 | BE042879 |
| C 36 | 573.4 | 53.5 | 925 | 9 | AI375908 |
| C 37 | 568 | 53.0 | 881 | 9 | AL541797 |
| C 38 | 567.2 | 52.9 | 652 | 9 | AW263746 |
| C 39 | 565.8 | 52.8 | 715 | 9 | AI887551 |
| C 40 | 564.4 | 52.6 | 727 | 10 | BE272013 |
| C 41 | 563.6 | 52.6 | 714 | 10 | BG032688 |
| C 42 | 561.2 | 52.4 | 877 | 10 | BE899344 |
| C 43 | 556.8 | 51.9 | 599 | 10 | BG818217 |
| C 44 | 549.6 | 51.3 | 649 | 9 | AW166916 |
| C 45 | 549.2 | 51.2 | 666 | 9 | AI888176 |

ALIGNMENTS

RESULT 1
AL572883/c
LOCUS
DEFINITION AL572883 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1034YC17 3 prime, mRNA sequence.
ACCESSION AL572883
VERSION AL572883.1 GI:12931579
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1018)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..1018
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1034YC17"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 199 a 306 c 331 g 172 t 10 others
ORIGIN

Db 638 GTGTGTCGAGGATCGCCAGCAGCTCTCCCGGCTGGCTACACCTCCACCTGTAAGGCT 579
QY 466 cgaatctcgagaagaagtgtctctgccagctgccaccttctctgcccgtagccct 525
Db 578 CGATCCCTCGAGAGGAAGTGTCTCTGCCAGCTGCCACCTCTCTGCCCGTAGCCCT 519
QY 526 cactgtgggtgaagagcactgagtggtggtggaagagacacttctgcatgataaacagacc 585
Db 518 CAGCTGGGTGTGAAGGACCTGGAGTGTGGGAAGGACACTTCTGCATGATAACAGACC 459
QY 586 tgcgtccgagacaacagcagggctgggctgctgctccctagccagggctctgttt 645
Db 458 TGTGTCGAGAGACAACCGAGAGGCTGGGCTGTCTGCTACCGCAGGCGGTCTGTCT 399
QY 646 gctgatcgccgacactgctgctctgctgcttccgctgcgcagcaggggtaccagtggt 705
Db 398 GCTGATCGGGCCACCTGCTCTGCTGCTGCTNTCCCTGCGCAGCCAGGGGTACCAAGTGT 339
QY 706 ttgctgagggagggcccgctgctgctgctgctgctgctgctgctgctgctgctgctg 765
Db 338 TTGCGCAGGAGGAGCCCGGCTGGGAGCGCCCTTTGAGGGAGCCAGCCCTTGAGACAGCTG 279
QY 766 ctgtgagagacactgaagactctgagcctcgagccctggagccctcgaggggtgcctc 825
Db 278 CTGTGAGGAGACAGTACTGAAGACTGTGACGCTCTCGGAGCCCTCGGAGGGGTGCCCTC 219
QY 826 tgcctgagggctcctgacactcctccctaccacaaattctcctggagccctattctgagct 885
Db 218 TGTGTCAGGCTCTCTGACACTCTCCCTAACCAATCTCCCTGAGCCCTTCTGAGCT 159
QY 886 ccccatcaccatggaggtgggctgctgctgctgctgctgctgctgctgctgctgctgctg 945
Db 158 CCCCATCACCATGGAGGTGGGCTGCTCAATCTAAGG-CTTTCCCTGTGAGAAGGGGTG 100
QY 946 agcacaagcc-cattacaagctgcatccctccctccctccctccctccctccctccctcc 1004
Db 99 TGGCAAAAGCCACATTACAAAGCTGCCATCCCTCCCGTTTTCAGTGGACCTGTGGCCAG 40
QY 1005 gtcttttccctatccacaggggtgttt 1032
Db 39 GTGCTTTTCCCTATCCACNAGGTGTT 12

RESULT 3
AL542659
LOCUS AL542659 LRI_FL002_PL1 Homo sapiens cdna clone CS0DE014YB02 5 prime
DEFINITION , mRNA sequence.
ACCESSION AL542659
VERSION AL542659.1 GI:12874922
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 942)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
CONTACT Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source location/Qualifiers
1. 942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE014YB02"
/clone_lib="LRI_FL002_PL1"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com

BASE COUNT 165 a 311 c 287 g 176 t 3 others
ORIGIN

Query Match 80.3%; Score 860.8; DB 9; Length 942;
Best Local Similarity 97.7%; Pred. No. 1.8e-188;
Matches 901; Conservative 3; Mismatches 15; Indels 3; Gaps 3;

QY 1 gggccccaccagtgccctggtggaagggccacacacacacacacacacacacacacacacac 60
Db 14 GGGCCCCACCAGGTGCCCTGGATGGAGAGGCCCCAGCTCACCTCAGCTGCCAGACCCA 73
QY 61 caagccttgagagagatgcccctgtgataatgtaagcagcgtgtccctcctccgtatcc 120
Db 74 CAAGCCTTGAAGAGAGATGCCCCCTGTGATAATGTCAGCAGCTGTCCCTCCTCCGATACC 133
QY 121 tgcgtccaaactcagctctgggagtggtggtggtggtggtggtggtggtggtggtggtggtg 180
Db 134 TGTGTCCTGCAATCTACGCTGGGGAGTGGGGCTGCTGTCCAAATCCCAGAGGCTGTCTGTGTC 193
QY 181 tcggaccacacagcactgctgccccacagcgtacacacacacacacacacacacacacacac 240
Db 194 TCGGACCACACGACACTGCTGCCCCAGGGTACACGCTGTGTAGCTGAGGGGACAGTGTCTG 253
QY 241 cgaggaaagcagatcgctggctggagtggtggtggtggtggtggtggtggtggtggtggtg 300
Db 254 CGAGGAAGCGAGATCGTGGCTGGACTGGAGAAAGATGCTGCCCGCCGCGCTTCTCTATCC 313
QY 301 caccacagagacatcggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 360
Db 314 CACCCAGAGACATCGGGCTGTGACACACACACACACACACACACACACACACACCTGTCTG 373
QY 361 ccgagccaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 420
Db 374 CCGAGCCTGGGTGGAGCTGGGCTGTGTCAGTTGCCCATGCTGTGTGTGTGCGAGGAT 433
QY 421 cgcagcactgctgcccggctgggtggtggtggtggtggtggtggtggtggtggtggtggtg 480
Db 434 CGCCAGCAGCTGCTGCCGGCTGGCTACACCTGCAACCTGTAAGGCTGCGATCTCGGAGAG 493
QY 481 gaagtgtctctccagcctgcccacacacacacacacacacacacacacacacacacacac 540
Db 494 GAAGTGGTCTCTGCCAGCCTGCCACCTTCTTGGCCCCGTAGCCCCCTACCTCAGCTGGGTG 553
QY 541 gacgtgagtggtgggagagacacttctgcatgataaccagacacacacacacacacacac 600
Db 554 GACCTGGAGTGTGGGAGGAGACACTTCTGCCATGATACACAGACCTGCTGCGGAGACAA 613
QY 601 cgacagggctgggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 660
Db 614 CGACAGGCTGGGCTGTGTGCTTCCCTACCGCCAGGGGCTGTGTGTGTGTGTGTGTGTGTG 673
QY 661 tgcgtcctgctggtcttcgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 720
Db 674 TGCTGTCTGTGCTGTCCCTGCGCAGGAGGGGTACCAAGTGTGTGCGCAGGAGG-C 732
QY 721 ccgctggagcgcctcttgaggac 780
Db 733 CCGCGCTGGGAGCGCCCTTTTGGAGGACCCAGCCCTTGACACAGCTGCTGTGAGGGACAG 792
QY 781 ctgaagactctgcagccctgcctgctgctgctgctgctgctgctgctgctgctgctgctg 840
Db 793 CTGAAGACTCTGCAGCCCTCGGGAGCCCGCCACTCGGAGGGGTGCCCTCTGTCTGAGGCC 852
QY 841 agcactccctcacaacaaattctcctggaccctattctgagctcccatcaccatggg 900

```

Db 853 AGCACTCCCTPAMCAATTCCTCTGGA-CCCAATTCAGCTCCCATCACCATGGG 911
QY 901 agtggggcctcaatcaagc 922
Db 912 AAGTGGGG-CTCATCTAGGC 932

RESULT 4
AL574944/c 900 bp mRNA linear EST 16-FEB-2001
LOCUS AL574944 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0DI064YG14 3
DEFINITION prime, mRNA sequence.
ACCESSION AL574944
VERSION AL574944.1 GI:12935632
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 900)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI064YG14"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 171 a 269 c 290 g 153 t 17 others
ORIGIN

Query Match 76.3%; Score 817.8; DB 9; Length 900;
Best local similarity 95.7%; Pred. No. 1.5e-178;
Matches 863; Conservative 16; Mismatches 19; Indels 4; Gaps 4;

QY 80 tccctgtgataatgtcagcagctgtccc-tctccgatactctgccaaactcagctc 138
Db 900 YCCCTGTGATATGTGACAGAGCTGTCCCGTCCCATACCTCTGCTCCCACTCACTGCT 841
QY 139 ggggagtggggctgtctccaaatccccagaggtgtctgtcgtcggaccacagcagctgc 198
Db 840 GGGGAGTGGGGCTGCTGTCCAAATCCAGAGGCTGTCTGTCTCGGACCACCACTGTC 781
QY 199 tggccacagcgtacacagctgtgtagctgaggggagctgtcagcagaggaagagatcgtg 258
Db 780 TGCCCCCAGGGCTACAGCTGTGTAGCTGTAGGGGAGGTGTACAGGAGGAGGAGATCGTG 721
QY 259 gctggactgagaagatgcttgcctccgcggttcccttattccaccacagacatcggc 318
Db 720 GCTGGACTGGAGAGATGCTCTG-CCGCCGGCTTCTTATATCCCAACCCAGAGACATCGC 662
QY 319 tgtgaccagcacacagctgcccggtggcggaacctgtgcccagagccaggggtggagc 378
Db 661 TGTGACCAGCACACAGCTGCCCGTGGGCGAGACCTGCTGCCCGAGGCTGTGGTGGGAGC 602
QY 379 tgggacctgtgccagttgccccatgctgtgtgtcgtcagagatgccagcagctgtgcgcc 438

```

```

Db 601 TGGGCTGTCTGCCAGTTGCCCATGCTGTGTGTCGAGATCGCCAGCACTGCTGCCCG 542
QY 439 gctggctacacctgcaacgtgaaggtcctgctcgcagagaagaagtggtctctgcccag 498
Db 541 GCTGGCTACACCTGCAACGTGAAGGCTCGATCTCTCGAGAGGAGNAGTGTCTCTGCCCAG 482
QY 499 cctggcaccttctcctggcccttagccctcactgctggtggtgaagacgtggtggtggtgaa 558
Db 481 CTTGCCACCTTCTCTGGCCCGCTAGCCCTCAGCTGGGTGTGAAGGACGTGGAGTGTGGGAA 422
QY 559 ggcacattctgcatgataaccagacgtgctgcagagacaacagagaggtggcctgc 618
Db 421 GRAMACTTCTGCCATATATACACAGCTGTGTCGAGACCAACCCAGAGGGCTGGGCTGC 362
QY 619 tgcctctagcccgagggcgtctgtgtgtgctgctgctgctgctgctgctgctgctgctt 678
Db 361 TGTCCTTACCGCCAGGGCGTCTGTGTGCTGTATCGGGGCCACATGCTCTCTGCTGGCTNT 302
QY 679 cgtgcgcacgcaggggtaccaggtgttgcgcagggagggccgcgcgtggagcgcct 738
Db 301 CCSTGGCGCAGCCAGGGGTACCAAGTGTTCGCGCAGGAGSSCCSGSGCTGGGACGSCCT 242
QY 739 ttgagggaaccagccttgagacagctgctgtgagggagcagctactgaagactctgcagccc 798
Db 241 TTGAGGAGACCCAGCGCTTGAGACAGCTGCTGTGAGGGAGACGTACTGAAGACTCTGCAGCCC 182
QY 799 tcgggacccactcggaggtgctcctcctgctcagggctcctcagcctcctcctcctaacca 858
Db 181 TCGGGACCCCACTCGGAGGGTGCCTCTGTCTAGGCTCTCCCTAGCAGCTCTCCSTFACCA 122
QY 859 aattcctcctggacccattctgagctcccatcaccatcaccatggaggtggggtcccaateta 918
Db 121 AATTCTCCCTGGACCCCATTTCTGAGCTCCCTCCCTCACCATGGAGGTGGGGCMTCAATCTA 62
QY 919 aggccttctcctgtcagaaggggttgagcaaaagcc-cattacaagctgcctccctcct 977
Db 61 AGG-CCCTTCCCTGTGAGAGGGGGTGTGGCAAAAGCCACATTAAAGCTCCCATCCCG 3
QY 978 cc 979
Db 2 CS 1

RESULT 5
AL576991/c 887 bp mRNA linear EST 16-FEB-2001
LOCUS AL576991 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0DI081YC03 3
DEFINITION prime, mRNA sequence.
ACCESSION AL576991
VERSION AL576991.1 GI:12939682
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI081YC03"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end

```

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com>

160 a 277 c 288 q 151 t 11 others

BASE COUNT

| Query Match | 71.4% | Score 765; | DB 9; | Length 887; |
|-----------------------|--|---------------------|-----------|-------------|
| Best Local Similarity | 94.3%; | Pred. No. 2.3e-166; | | |
| Matches 836; | Conservative 5; | Mismatches 41; | Indels 5; | Gaps 5; |
| Qy 131 | tcaagtgtgggaagtgggctgtcgaatccagaggctgtgtctgtcgcgaccacc | 190 | | |
| Db 887 | TCAAGTGTGGGAAGT-GGGCTGTGTCCAAATCCAGAGGCTTKTGTCTGCTCGGACACC | 829 | | |
| Qy 191 | agcaactgtccccagcgaTacaactgtgtagctgtgagggcagtgctcagcgaggaagcg | 250 | | |
| Db 828 | AGMACGGCTGCCCCAGGGGCAACAGTGTAKTGTAGCTAGGGGCGAGTGTCTAGCGAGGACGG | -769 | | |
| Qy 251 | agatcgtggctggactggagaagatg-cctgccgcgcgcggttccctatccaccaccaga | 309 | | |
| Db 768 | AGATCGTGGCTGGACTGAGAAAGATGCCCTGCCCGCGGGCTYCTTTATCCACCCACAGA | 709 | | |
| Qy 310 | gacatcggctgtgaccagcacac-cagctgcgcgcggtggcggaacctgtctgcccgagcca | 368 | | |
| Db 708 | GACATCGGCTGTGACCAAGCACACAGCTGCCCGGTGGGCGAGAGCTCTCTGCCGAGGCT | 649 | | |
| Qy 369 | gggtgggaagctggccctgtcgcagttgcccatctctgtctgtcgcagagatgccagca | 428 | | |
| Db 648 | GGGTGGGAGCTTGGGCGGGCTGCCAGTGCCTCCCATGTCTGTGTCTGCGAGGATCGCCAGMA | 589 | | |
| Qy 429 | ctgtgccgcgctggctacacctgcacacgtgaaagctcgatcctgcgagaaggaagtgt | 488 | | |
| Db 598 | CTGTGCCCGGCTGGCAACACCTGCAACGTGAAGCTCGATCCTGCGAGAGGAAGTGGT | 529 | | |
| Qy 489 | ctctgccagctgccacctcctggccgctagccctcaactgtgtgtgaaggacgtgga | 548 | | |
| Db 528 | CTCTGCCAGAGCTTGCCACTGCCCTGGCCCGGTAGCCCTCACGTGGGTGTGAAGAGCTGGA | 469 | | |
| Qy 549 | gtgtgggaagacaactctccatgataaacagacctgtgcgcagacacaccacacagg | 608 | | |
| Db 468 | GTGTGGGGAAGACACTTCTGCCATGATPACACAGACACMGCTGCCGAGAAACACCGACAGGG | 409 | | |
| Qy 609 | ctgggcctgtctccctacgcgcgcgcgcgcgtgtgtgtgtctgacgcgcacctgtctcc | 668 | | |
| Db 408 | CTGGGCCNGCTGTCCCTACCGCCAGGGGCTGTGTGTCTGATCGCGCCACNGCTGTCC | 349 | | |
| Qy 669 | tgtgtgcttcgcgtgcgcacgcaggggtaccgaagtgtttgcgcagggagggccccgcgtg | 728 | | |
| Db 348 | GGCTGGCTTCGG-TGCGCAGCAGCGGGGTACCAAGTGTGTGGCGCAGGGAGGCCCGCGTG | 290 | | |
| Qy 729 | ggacgcctcttgagggaaccagccttgagacagctgtgtgaggacagtactgaagac | 788 | | |
| Db 289 | GGACGCCCTTTGAGGGACCCAGCCTTGAGACAGCTGTGTGTGAGGGGACAGAACTGAAGAC | 230 | | |
| Qy 789 | tctgcagccctcgggaaccaccactcggaggggtgcctctgtcagggcctccctagcacctc | 848 | | |
| Db 229 | TCTGCAGCCCTCGGGACCCCACTCGGAGGGTGCCTCTNGCTCAGGCTCCCTTAGCACCTC | 170 | | |
| Qy 849 | cccttaaccaaatctcctcggaccccaattctgagctcccccatcacctggaggtgggg | 908 | | |
| Db 169 | TCCCTTAACCAAAATTCCTCGTGAGCCCAATCTGAGCTCCCATCACCTTGGGAGGTGGGG | 110 | | |
| Qy 909 | cctcaatctaaggccctcctcgtcagaagggggttgaggcaaaagcc-cattacaagct | 967 | | |
| Db 109 | CCTCAATCTAAGGCCCTTCCCTGTCAAGAGGGGGTNTGTGGGAAAAGCCACATAACAAGT | 50 | | |
| Qy 968 | gccatccctcctccctttcaatgaacctgtggccagctgtctttccct | 1016 | | |

| | | | |
|------------|-----|---|---|
| Db | 49 | GCATCCCCCTCCCGTTTCAGTGGACCTGTGGCCAGGTGCTTTTCCTT | 1 |
| RESULT | 6 | | |
| LOCUS | 182 | | |
| DEFINITION | 182 | | |
| ACCESSION | 182 | | |
| VERSION | 182 | | |
| KEYWORDS | 182 | | |
| SOURCE | 182 | | |
| ORGANISM | 182 | | |
| REFERENCE | 182 | | |
| AUTHORS | 182 | | |
| TITLE | 182 | | |
| JOURNAL | 182 | | |
| COMMENT | 182 | | |

| | Query Match | 70.9% | Score 760.4; | DB 10; | Length 810; |
|----|-----------------------|--|---------------------|-----------|-------------|
| | Best Local Similarity | 97.9%; | Pred. No. 2.6e-165; | | |
| | Matches 781; | Conservative 0; | Mismatches 16; | Indels 1; | Gaps 1 |
| Qy | 10 | caggtgcctgattggagaagcccgacgtcacctcagctgccagaccacacaagccttg | 69 | | |
| | | | | | |
| | | | | | |
| | | | | | |
| Dd | 2 | CAGGTGCCTCGATGGAGAAGGCCCCAGCTCACCTCAGCTGCCAGACCCACAAGCCTTG | 61 | | |
| | | | | | |
| Qy | 70 | aagagagatgtcccctgtgataatgtcagcagtcgtcccctctccgatcactgtgcctcaa | 129 | | |
| | | | | | |
| | | | | | |
| | | | | | |
| Dd | 62 | AAGAGAGATGTCCTCTGTGAATAATCTCAGCAGCTGTCCCTCTCCGATACCTGCTGCCAA | 121 | | |
| | | | | | |
| Qy | 130 | ctcagctctggggagtgaggctgctgtcccaatcccagaggctgtctgtctgcctcgagaccac | 189 | | |
| | | | | | |
| | | | | | |
| Dd | 122 | CTCAGTCTGGGAGTGGGGCTGCTGTCCAATCCAGAGGCTGTCTGCTCGGACCAC | 181 | | |
| Qy | 190 | cagcactgtctcccccaagcgtatacagtgtgtagctgtgaggggcagtgctcagcgaggaagc | 249 | | |
| | | | | | |
| Dd | 182 | CAGCACTGCTCCCCCAGGGGTACACGTGTGTAGCTGAGGGGCGCATGCTCAGGGAGGAAGC | 241 | | |
| Qy | 250 | qagatctgttgcctgactgtagaagatgcctgcgcgcgcgcggttctcttatccacccccaga | 309 | | |

```

|||||
Db 242 GAGATCGTGGCTGGACTGGAGAGATGCTGCCCGCCGGGCTTCCTATCCACCCAGCA 301
Qy 310 gacatcgctgtgacacgacacacagctgcccgggtggcggaacactgtgcccagcag 369
Db 302 GACATCGGCTGTGACACGACACAGCTGCGCGGTGGGCGACACCTGCTGCCCGAGCTG 361
Qy 370 ggtggagctgggctgtgcccagttgcccacatgtgtgtgtgtgtgtgtgtgtgtgtgt 429
Db 362 GGTGGAGCTGGGCTGTGCTGCCAGTTGCCATGCTGTGTGTGTGTGTGTGTGTGTGTGT 421
Qy 430 tctctccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 489
Db 422 TGTGCTCCCGGCTGGGTACACCTGCAACGTGAAGGCTCGATCTGCGAGAGAAAGTGTTC 481
Qy 490 tctgcccagctgtccacacttctgtcccgtagccctcagctgtgtgtgtgtgtgtgtgtgt 549
Db 482 TCTGCCAGCTGCCACCTTCTGTGCCGTGACCCCTACGCTGCGGTGTGTGTGTGTGTGT 541
Qy 550 tgtggtggaagacacttctgtccatgataaccagacactgtgtgtgtgtgtgtgtgtgtgt 609
Db 542 TGTGGGGAAGGACACTTCTGTGCCATGATACACAGACTGTGTGTGTGTGTGTGTGTGTGT 601
Qy 610 tgggctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 669
Db 602 TGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
Qy 670 gctggttccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 729
Db 662 GCTGGCTTCCGCTGGCAGCCAGGGGTACCAAGTGTATGCGGAGGAGGCCCGG-GCTGG 720
Qy 730 gacgcccctttagggagccagccttgagacagctgtgtgtgtgtgtgtgtgtgtgtgtgt 789
Db 721 GACGCCCTTTGAGGGACCCAGGCTTGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy 790 ctgacgacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 807
Db 781 CTGACGCCCTCGGGAGCCC 798

RESULT 7
AL562633/c
LOCUS AL562633 LTI_NFL003_NBC3 908 bp mRNA linear EST 16-FEB-2001
DEFINITION prime, mRNA sequence.
ACCESSION AL562633
VERSION AL562633.1 GI:12911246
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 908)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
location/Qualifiers
1..908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC015YA01"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
```

```

vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com*
BASE COUNT 153 a 292 c 296 g 157 t 10 others
ORIGIN
Query Match 70.6%; Score 757; DB 9; Length 908;
Best Local Similarity 93.3%; Pred. No. 1.6e-164;
Matches 847; Conservative 6; Mismatches 49; Indels 6; Gaps 6;
Qy 98 gagctgttctctccctccgataacctgtg-ccaaactcaagctgtggggagtggggctgtgt 156
Db 908 GGGGCTGTCTCTCTCCGATACCCCTGCCCAATACACATCTCTGGGAGTGGGGCGCTGT 849
Qy 157 ccaa-tcccagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 215
Db 848 CCAATCCAGAGGCTGTGTCTGCTGCGACACACAGACGCTGCCCGGCAACAC 789
Qy 216 gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 275
Db 788 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
Qy 276 gctgtccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 335
Db 728 GCYTSGCCCGGGCTTCCTTATCCACCAAGARACATCGCTGTGTACACACACACAG 669
Qy 336 ctgcccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 395
Db 668 CTGCCCGTGGGGCAGACCTGTGTGCCGAGCTGGGTGGGAGCTGGGCTGTGTGCCAGTT 609
Qy 396 gcccagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 454
Db 608 GCCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 549
Qy 455 acgtgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 514
Db 548 ACCTGAAGGCTCGATCTCTCGAGAGAGAAAGTGTCTCTGCCAGCTGTGCACCTNCTGG 489
Qy 515 cccgtagcctcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 574
Db 488 CCCGTAGCCCTCACCCTGGGTGTGAAGAGCTGTGGAGAGTGTGGGAGAGACACTTCTT 429
Qy 575 ataacagacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 634
Db 428 ATAACAGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 369
Qy 635 gcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 694
Db 368 CGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 309
Qy 695 gtacaaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 754
Db 308 GTACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 249
Qy 755 tgagacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 814
Db 248 TGAGAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189
Qy 815 aggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 874
Db 188 AGGGTGGCTCTGCTCAGGCGCTCCCTAGCACCTCTCCCTTAACCAAAATTCCTCCTGG 129
Qy 875 catctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 934
Db 128 CATCTGAGCTCCCATCACCATGGGAGTGGGGGCTCAATCTAAGG-CTTTCCTGTCA 70
Qy 935 gagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 992
Db 69 GAAGGGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10
```



```
BASE COUNT      156 a      262 c      266 g      140 t      11 others
ORIGIN

Query Match      68.5%; Score 734.8; DB:9; Length 835;
Best Local Similarity 96.5%; Pred. No. 2.1e-159;
Matches 782; Conservative 10; Mismatches 14; Indels 4; Gaps 4;

QY 147 gggctgtcccaatccagagctgtctgtctgctgagccaccagcactgctgcccaca 206
DB 806 GGGCTGTGTCATCCAGAG-CGTGTCTGTCTGCTGGACCACAGCAGCTGTGCTGCCCA 748
QY 207 gcgatacaagtgtagtgaggggagtgtagtgtagcaggaagcagagatggtgagct 266
DB 747 GGGCTACACGTGTGTGCTGAGGGGAGTGTGAGCGAGGAGGAGAGTGTGTGCTGGACT 688
QY 267 ggaagagatcctgcccgcgcgggttccttatcccccagagacatggctgtgacca 326
DB 687 GGAGAGATGCCGTG-CGCGGGGCTTCTTATCCACCCAGAGACATCGGCTGTGACCA 629
QY 327 gcacaccagctgcccgtggcggaacctgtgctgcccagccaggtggagctggccctg 386
DB 628 GCACACCAGCTGCCGGTGGGGGAGACCTGCTGCCGAGGCTGGTGGGAGCTGGGCTG 569
QY 387 ctgccagttgccccatgtgtgtgtgctgagagatccagcactgtgcccggctggcta 446
DB 568 CTGCMAGTTGCCCATGCTGTGTGTGCTGAGGATGCGCAGCAGTGTGTGCCGGCTGCTA 509
QY 447 cactgcaacgtgaagctgactgctgaggaaggaagtggtctctgcccagctgcccac 506
DB 508 CACTGTCAACGTGAAGCTCGATCTGCGAAGGAAGTGTGTCTGTGCCAGGCTGCCAC 449
QY 507 ctctctggccctgagcctcactgagtgagagcagtgagtgagtgagtgagtgagc 566
DB 448 ITTCTGGCCCGTAGYCTCCTGCTGGGTGTGAAGGAGCTGGAGTGTGGGAAGAHACTT 389
QY 567 ctgcatgataaacagacctgtgcccagagacacacagggctggcctgtgctcccta 626
DB 388 CTGCCATGATAACAGACCTGTGCTGCGAGACACCGACAGGGCTGGGCTGTGCTCCTA 329
QY 627 cgcacagggcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 686
DB 328 CCGCAGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 270
QY 687 acgaggggtaccagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 746
DB 269 AGCCAGGGGTACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 210
QY 747 cccagccttgagacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 806
DB 209 CCCAGCCTTGAGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 150
QY 807 ccactcagaggggtgacctgtgctcagcctcctcagcctcctcagcctcctcagcct 866
DB 149 CTACTCGAGGGGTGCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 90
QY 867 ctgagccccattctgagctccccatccatggaggtgggaggtgggaggtgggaggtgg 926
DB 89 CTGGACCCCATCTGTAGCTCCCATCACCATGGGAGGTGGGAGGTGGGAGGTGGGAGTCA 31
QY 927 cccctgcagaaggggtgtgaggaagcc 956
DB 30 CCCTGTGAGAAGGGGKGTGTGGCAAAAGCC 1

RESULT 10
BG763617
LOCUS      602735831F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4861255 5',
DEFINITION mRNA sequence.
ACCESSION BG763617
VERSION    BG763617.1 GI:14074270
KEYWORDS   EST.
```


[illegible]

QY 718

718 db

REFERENCE
AUTHORS

JOURNAL
COMMENT

FEATURES SOURCE

01000

BASE COUNT

ORIGIN

Query Ma

Best Local Matches

Qy 138

Db 894

Qy 197

834 Db

QY 255 cgtgctggac--tggagaagatgctcccgccgctgcttcttcttccacccacccagagac 312
|||||
Db 774 COTGCTGGACCTGGAGATAATGCTGCCCGCGGCTTCTGATCCACCCAGAGAC 715
QY 313 atcgctgtgaccagacacacagctgcccgtggcgga-acctctgccagccaggg 371
|||||
Db 714 ATCGCTGTGACACACACACAGCTGCCGCTGAGGGGAGGAGCTGCTGCCGAGCTGGG 655
QY 372 tggagctgggctg-ctgcaagttgccccatgctgtgtgctgagagatgcccagac- 429
|||||
Db 654 TGGAGCTGGGCTGCTGCCAGTTGCCCATGCTGTGTGCTGCCAGGATGCCAGCACA 595
QY 430 tctgccccgc-tggctacacctgcaacgtgaagctgctgctgctgagagaagtgt 488
|||||
Db 594 TCTGCCCGGCATGCTACACCTGCAACGTGAAGCTCGATCTGCGAGAAGAGTGT 535
QY 489 ctctccacagctgacacttctgcccgtgagccctgacccctgagtgagagctgga 548
|||||
Db 534 CTCTGCCAGCTGCCACCTCTCTGCCGCTGAGCCCTACGTGGGTGAAGACCTGGA 475
QY 549 gtgtgggaagacacttctgcatgataacacagacactgctgcccagacacccagag 608
|||||
Db 474 GTGTGGGGAAGACACTTCTGCCATGATAACACAGACCTCTGCCGAGACACCGACAGG 415
QY 609 ctgggctgctgctcctacgccccaggggctgtgtgtgtgtgtgtgtgtgtgtgtgt 668
|||||
Db 414 CTGGGCTCTGCTCTCTACCGCCAGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 355
QY 669 tctggtctccgctgcaagcaggggtaccagttgttggcagggagggcccgctg 728
|||||
Db 354 TGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 295
QY 729 ggaagccctttagagagacccagcttgagacagctgtgtgagggacagactggaagac 788
|||||
Db 294 GNAAGCCCTTGTAGGAGCCAGCTCTGAGACAGCTGTGTGAGGACAGACTGAGAGAC 235
QY 789 tctgagccctgggagccacactggaggggtgcccctctgctcagggccctcctagacctc 848
|||||
Db 234 TCTGACGCTCTGGGACCCACCTCGAGACAGCTGTGTGAGGACAGACTGAGAGCT 175
QY 849 cccctaacaattctcctggagccacattctgagctcc-ccatcaccatggaggtgg 907
|||||
Db 174 CCCTTAACCAATTCCTCTGGACCCCACTCTGAGCTGCTGAGCTGCTGAGGAGTGGG 115
QY 908 gctcaatcaagcccttccctgctcagaggggtgaggaagcc-cattacaagc 966
|||||
Db 114 GCCTCAATCTAAGG-CTTCCCTGTGAGAGGGGGTGTGGCAAGGCCACATTACAAGC 56
QY 967 tgcataccctccctcccttccagtgagcc 994
|||||
Db 55 TGCCATCCCTCCCTCCCTTTCAGTGGACC 28

RESULT 14
BG830299
LOCUS 602764696F1 NIH_MGC_42 Homo sapiens CDNA clone IMAGE:4906991 5',
DEFINITION mRNA sequence.
ACCESSION BG830299
VERSION BG830299.1 GI:14177886
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 751)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1810 row: o column: 24
High quality sequence stop: 748.
FEATURES
Location/Qualifiers
source
1..751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4906991"
/clone_lib="NIH_MGC_42"
/tissue_type="epitheloid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
BASE COUNT 125 a 249 c 227 g 150 t
ORIGIN

Query Match 62.8%; Score 673.2; DB 10; Length 751;
Best Local Similarity 96.5%; Pred. No. 3.4e-145;
Matches 720; Conservative 0; Mismatches 23; Indels 3; Gaps 3;
QY 286 cgcggtcttctatccacccacagagacatcggtgtgaccagcacacacagctgcccggtg 345
Db 2 CGGGCTTCTTATCCACCCACAGAGACATCGCTGTGACCCAGCACACAGCTGCCGGTG 61
QY 346 ggcggaaactgtctgcccagcaggggtggagctgggctgtgcccagttgcccagctgct 405
Db 62 GGGCAGACCTGTCTGCCGAGCCTGGGTGGAGGTGGGCCCTGTCTGCCAGTGTGCCCATGT 121
QY 406 gtctgtcagagatgcagcagctgctgcccgtgtgtacacctgcaacgtgaagct 465
Db 122 GTGTGTCTGCCAGAGATGCCAGACACTGTCTGCCGGTGGCTACACCTGCAAGCTGAAGCT 181
QY 466 cgtctcgcagaggaagtgtgtctgtcccagcctgcccacctctctctggtgcccctagccct 525
Db 182 CGATCTTGCAGAGGAAGTGGTCTCTGCCAGCCTGCCACCTTCTCTGGCCGCTAGCCCT 241
QY 526 cactggtgtgagagcagctggaggtggggaagagacacttctgccaatgataccagacc 585
Db 242 CACGTGGTGTGAAGGACGTGGAGTGTGGGGAAGGACACTTCTGCCATGATAACACAGACC 301
QY 586 tgctgcagacacacagcagggctgctgctgctcctacgcccagggcgctgtgtgt 645
Db 302 TGCTGCCAGACACACACAGAGGCTGGCCCTGTGTCTCTACCCAGAGGGCTGTGTGT 361
QY 646 gctgctggtgagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 705
Db 362 GCTGATCGGCGCCACTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 420
QY 706 ttgcgcagggagggcccgctggagcggcccttggaggaccacagccttgagacagctg 765
Db 421 TTGGCAGGGAGGGCCCGCTGGGACGCCCTTTTGGAGGACCCAGCCTTTGAGACAGCTG 480
QY 766 ctgtgagggagcagctgagagctgagcctcctgagccctcgagccactcgagggtgcccctc 825
Db 481 CTGTGAGGGAGACAGTACGTAAGACTCTGCAGCCCTCGGAGACCCCACTCGGAGGGTGCCTC 540
QY 826 tctcagggcctcccttagcacctccctcccttaacaaattctcctgagccctcttctgagct 885
Db 541 TGCTCAGGCTCTCTTAGCACCCTCCCTACCAATTTCTCCCTGGACCCCACTTCTGAGCT 600
QY 886 ccccatcaccatgggaggtgggctcaatcaatgaaggcccttccctgtcagaaggggtgtg 945



| Result No. | Score | Query | Length | DB | ID | Description |
|------------|-------|-------|---------|----|------------------|-------------------|
| 1 | 1072 | 100.0 | 2095 | 4 | US-08-991-862-16 | Sequence 16, Appl |
| 2 | 747.2 | 69.7 | 1779 | 1 | US-07-668-648-3 | Sequence 3, Appl |
| 3 | 747.2 | 69.7 | 1779 | 2 | US-08-429-998-3 | Sequence 3, Appl |
| 4 | 747.2 | 69.7 | 1779 | 2 | US-08-431-333-3 | Sequence 3, Appl |
| 5 | 747.2 | 69.7 | 1779 | 5 | PCR-US91-02321-3 | Sequence 3, Appl |
| 6 | 533.8 | 49.8 | 2137 | 4 | US-08-991-862-1 | Sequence 1, Appl |
| 7 | 469.2 | 43.8 | 1767 | 2 | US-07-668-648-1 | Sequence 1, Appl |
| 8 | 469.2 | 43.8 | 1767 | 2 | US-08-429-998-1 | Sequence 1, Appl |
| 9 | 469.2 | 43.8 | 1767 | 2 | US-08-431-333-1 | Sequence 1, Appl |
| 10 | 469.2 | 43.8 | 1767 | 5 | PCR-US91-02321-1 | Sequence 1, Appl |
| 11 | 448.4 | 41.8 | 1767 | 1 | US-07-668-648-5 | Sequence 5, Appl |
| 12 | 448.4 | 41.8 | 1767 | 2 | US-08-429-998-5 | Sequence 5, Appl |
| 13 | 448.4 | 41.8 | 1767 | 2 | US-08-431-333-5 | Sequence 5, Appl |
| 14 | 448.4 | 41.8 | 1767 | 5 | PCR-US91-02321-5 | Sequence 5, Appl |
| 15 | 153.2 | 14.3 | 539 | 1 | US-07-668-648-7 | Sequence 7, Appl |
| 16 | 153.2 | 14.3 | 539 | 2 | US-08-429-998-7 | Sequence 7, Appl |
| 17 | 153.2 | 14.3 | 539 | 2 | US-08-431-333-7 | Sequence 7, Appl |
| 18 | 153.2 | 14.3 | 539 | 5 | PCR-US91-02321-7 | Sequence 7, Appl |
| 19 | 64.6 | 6.0 | 341 | 1 | US-07-668-648-9 | Sequence 9, Appl |
| 20 | 64.6 | 6.0 | 341 | 2 | US-08-429-998-9 | Sequence 9, Appl |
| 21 | 64.6 | 6.0 | 341 | 2 | US-08-431-333-9 | Sequence 9, Appl |
| 22 | 64.6 | 6.0 | 341 | 5 | PCR-US91-02321-9 | Sequence 9, Appl |
| 23 | 37 | 3.5 | 20303 | 1 | US-08-370-975B-6 | Sequence 6, Appl |
| 24 | 37 | 3.5 | 26764 | 1 | US-08-370-975B-1 | Sequence 1, Appl |
| 25 | 36.2 | 3.4 | 4403765 | 4 | US-09-103-840A-2 | Sequence 2, Appl |
| 26 | 36 | 3.4 | 935 | 4 | US-08-822-038-73 | Sequence 73, Appl |
| 27 | 36 | 3.4 | 935 | 4 | US-08-479-285-73 | Sequence 73, Appl |

| | | | |
|----|------|---|------|
| Db | 1246 | AAGGAGACAGATGGTGGCTGGAGCATGCTGCCGACAGCACTCTGCTC | 1300 |
| Qy | 301 | cacccagagacatcggtgtgaccagcacacccagctgccccggtggcggaacctgtgcg | 360 |
| Db | 1306 | CAACATGGAGATATTGGTTGTGACCAGCATACACAGTGCCTGCCAGTGGGCAACATGCTGC | 1365 |
| Qy | 361 | ccgagcgaaggtggagctggcctgtccagttgccccatgctgtgtgtgcgaggat | 420 |
| Db | 1366 | CCAAGCTGAAGGGAAGTTGGGCCCTGCTCCAGTTGCCCATGCTGTGTGCTGTGAAGAC | 1425 |
| Qy | 421 | cgcagcactgtgccccggtggtacacctcaacgtgaaaggtctcgatctctgcgagaag | 480 |
| Db | 1426 | CGGAGCACTGTTGCCCGCTGGGTACACTGCACAGCTGAAGCGCAGAACTGTGAGAAG | 1485 |
| Qy | 481 | gaagtgtctctgccagcctgcceacttctggccgtagccctacgtgggtgtgaag | 540 |
| Db | 1486 | GATGCAGGCTCTGTCCAGCTTCCATGAGCCCTTTGGCTCTAAGGTGG-----G | 1539 |
| Qy | 541 | gacgtggagtggtgggaagacactctgtccatgataaccagacctgctgccgagacaac | 600 |
| Db | 1540 | AATGTGGAATGTGGTCCGGACATTTGSCCATGATACACAGTCCCTGTTGTAAGACAGC | 1599 |
| Qy | 601 | cgacagggtggcctgtgtccctacgccccagggcgctgtgtgtctgatcggcgccac | 660 |
| Db | 1600 | CAAGGAGGCTGGGCCCTGCTCCCTATGTATAAGGGTGTCTGCTGTAGAGATGGACGTCA | 1659 |
| Qy | 661 | tgtctctctgtggtcttcgctgcgcacgcaggggtaccgaagtgtttgcgcagggagggc | 720 |
| Db | 1660 | TGTTGTCCCATTTGGCTTCCACTGTTACGCCAGGGGNAACCAAGTGTTCGGGAAGAGACC | 1719 |
| Qy | 721 | ccgcgtgggacgcccccttgagggacccagccttgagacagctgtg | 768 |
| Db | 1720 | CCTCGCTGGACATACTTTTGAAGGATCCAGCCCCAAGACCGCTACTG | 1767 |

RECEIVED

US-08-429-998-1
 ; Sequence 1, Application US/08429998
 ; Patent No. 5885961
 ; GENERAL INFORMATION:
 ; APPLICANT: Shoyab, Mohammed
 ; APPLICANT: Plowman, Gregory D.
 ; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
 ; TITLE OF INVENTION: MODULATING PROTEINS
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/429,998
 ; FILING DATE: 27-APR-1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/668,648
 ; FILING DATE: 13-MAR-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mistrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 5624-161-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212)790-9090
 ; TELEFAX: (212) 869-9741
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:

```

Db   2042 tcccgagccacgagtgtctgtgactgttctgtg 2076
||||| ||||| || | ||| ||||
RESULT      7
US-07-668-648-1
; Sequence 1, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plozman, Gregory D.
; TITLE OF INVENTION: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: EPIPELINS: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/668,648
; FILING DATE: 19910819
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-161-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1767 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1767
; US-07-668-648-1

Query Match          43.8%; Score 469.2; DB 1; Length 1767;
Best Local Similarity 76.7%; Pred. No. 8.6e-107;
Matches 589; Conservative 0; Mismatches 173; Indels 6; Gaps 1;

Qy    1  gggccccaccaggctgccttgatgagagcccagctcacctcagctgccagaccga 60
       ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1006 GGAGTCCTTCCAGGTACCTTGATGAAAGGTACGGCCTCCCTCAGGCTGCCAGACCCA 1065
       ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy    61  caagcctttaagagagatgtccctctgtataatgtcagcagctgtccctctccgatcc 120
       ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1066 CAGATCTTTGAAGAATGATGTCCCCTGTGATGACTTCAGTAGTGTCTTAACAATAACC 1125
       ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy    121  tgcgtccaactcaagctctgggaagtgggctgctgtccaaatccccagagagctgtctgtgc 180
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1126 TCGCTGCAGACTCAGTCTTTGGGGAGCTGGGGCTGCTGTCCCATCCAGAGGCTGTCTGCTGC 1185
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy    181  tcggaaccaccagcactgctgccccagcagatcacacgtgttagctgaggggcagtgctcag 240
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1186 TTAGACCACCAAGCATTTGTCCTCCCTCAGGGTTCAAATGTATGGATGAGGGGTACTGTGAC 1245
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy    241  cgaggaagcgsagatcgtgtgctggagtggagaagatgacctgcccccgcgcggttccttatcc 300
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1767
US-08-429-998-1

Query Match 43.8%; Score 469.2; DB 2; Length 1767;
Best Local Similarity 76.7%; Pred. No. 8.6e-107;
Matches 589; Conservative 0; Mismatches 173; Indels 6; Gaps 1;
QY 1 gggcccccacacaggtgcccctgagtgagagagggcccccagctcacctcagcctgccagaccac 60
DB 1006 GGAGTCTCTTCAGTACCTGGATGAAAAGGTCAAGGCTCCCTCAGCTGCCAGACCCA 1065
QY 61 caagccttgaagagagatgccctgtgataatgtcagcagctgtccctcctccgatacc 120
DB 1066 CAGATCTTGAAGATGATGCTCCCTGTGATGACTTTCAGTGTCTTCAACAATACC 1125
QY 121 tgcctgcaactcacgtctggggagtgggctgtgtcccaatcccagaggtgtgtctgtgc 180
DB 1126 TGCTGCAGACTCAGTCTTGGGGACTGGGGCTGTCTCCCAATGCTGTCTTCAACAATACC 1125
QY 181 tcggaccacacagcactgtgtcccccagcgcatacacgtgtgtagctgagggcagtgctgc 240
DB 1186 TTAGACCACACAGCATTTGCTCCCTCAGGCTTTCAAAATGATGATGAGGGGTACTGTGTCAG 1245
QY 241 cgaggaagcagagatcgtgtgagtgagagagatgcctgcgcgcgcgcgcgcgcgcgcgc 300
DB 1246 AAGGAGACAGAAATGCTGCTGGCTGGCTGGAGAGATGCTGTCCGCCAGACAACTCTGTCTC 1305
QY 301 cacccccagagacatcggtgtgacacagcagcagcagcagcagcagcagcagcagcagcagc 360
DB 1306 CAACATGGAGATATTGTTGTGACGACATACACAGCTGCCCCAGTAGGGCAACATGCTGTC 1365
QY 361 ccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420
DB 1366 CCAAGCCTGAAGGAATGTTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1425
QY 421 cgccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
DB 1426 CGGAGGACTGTTCCCGGCTGGGTACACTGCAACGTGAAGGCGAGAACCTGTGAGAG 1485
QY 481 gaagtgtctctgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540
DB 1486 GATCAGGCTCTGTCCAGCCTTCCATGACCTGACCTTTGGCTCTAAGGTTGG-----G 1539
QY 541 gacgtgagtgaggagagacactctctccatataacacagcagcagcagcagcagcagcagc 600
DB 1540 AATGTGAATGTGTGGCGGACATTTCTGCCATGATAACCACTGCTGTGTAAAGACAGC 1599
QY 601 cgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 660
DB 1600 CAAGGAGCTGGGCTGCTGCTCCATATGTAAGGGTGTCTGCTGTGAGAGATGGAGTCTAC 1659
QY 661 tgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 720
DB 1660 TGTTGCTCCATGCTTCCATGTTTACGCAAGGGAACCAAGTGTGTCGGAAGAAGACC 1719
QY 721 ccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 768
DB 1720 CCTGCTGGGACATACCTTTTGGGGATCCAGCCCCAGACCGCTACTG 1767

RESULT 9

US-08-431-333-1

; Sequence 1, Application US/08431333

; Patent No. 5965723

GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
APPLICANT: Plowman, Gregory D.
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
TITLE OF INVENTION: MODULATING PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,333
FILING DATE: 27-APR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/668,648
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 869-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1767
US-08-431-333-1

Query Match 43.8%; Score 469.2; DB 2; Length 1767;
Best Local Similarity 76.7%; Pred. No. 8.6e-107;
Matches 589; Conservative 0; Mismatches 173; Indels 6; Gaps 1;
QY 1 gggcccccacacaggtgcccctgagtgagagagggcccccagcctcacctcagcctgccagaccac 60
DB 1006 GGAGTCTCTTCAGTACCTGGATGAAAAGGTCAAGGCTCCCTCAGCTGCCAGACCCA 1065
QY 61 caagccttgaagagagatgccctgtgataatgtcagcagctgtccctcctccgatacc 120
DB 1066 CAGATCTTGAAGATGATGCTCCCTGTGATGACTTTCAGTGTCTTCAACAATACC 1125
QY 121 tgcctgcaactcacgtctggggagtgggctgtgtcccaatcccagaggtgtgtctgtgc 180
DB 1126 TGCTGCAGACTCAGTCTTGGGGACTGGGGCTGTCTCCCAATGCTGTCTTCAACAATACC 1125
QY 181 tcggaccacacagcactgtgtcccccagcgcatacacgtgtgtagctgagggcagtgctgc 240
DB 1186 TTAGACCACACAGCATTTGCTCCCTCAGGCTTTCAAAATGATGATGAGGGGTACTGTGTCAG 1245
QY 241 cgaggaagcagagatcgtgtgagtgagagagatgcctgcgcgcgcgcgcgcgcgcgcgc 300
DB 1246 AAGGAGACAGAAATGCTGCTGGCTGGCTGGAGAGATGCTGTCCGCCAGACAACTCTGTCTC 1305
QY 301 cacccccagagacatcggtgtgacacagcagcagcagcagcagcagcagcagcagcagcagc 360
DB 1306 CAACATGGAGATATTGTTGTGACGACATACACAGCTGCCCCAGTAGGGCAACATGCTGTC 1365

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 5, 2002, 16:30:39 ; Search time 3362.64 Seconds
(without alignments)
6671.320 Million cell updates/sec

Title: US-09-824-647-16_COPY_1024_2095
Perfect score: 1072
Sequence: 1 gggccaccaggcgccctg.....ataaagttgtcaacttttt 1072

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB ID | Description |
|------------|----------|-------|-------|--------|--------------|-------------------------------------|
| 1 | ARI75456 | 100.0 | 100.0 | 425 | US 6309826-A | 16 30-OCT-2001; Location/Qualifiers |

| | | | | | | | |
|----|-------|-------|--------|----|--------------|----|---------------------|
| 1 | 1072 | 100.0 | 2095 | 6 | ARI75456 | 6 | ARI75456 |
| 2 | 1072 | 100.0 | 2095 | 9 | HUMGRANUL | 9 | HUMGRANUL |
| 3 | 992.6 | 92.6 | 2132 | 9 | BC010577 | 9 | BC010577 Homo sapi |
| 4 | 992.6 | 92.6 | 2152 | 9 | HSEPT1 | 9 | HSEPT1 H. sapiens m |
| 5 | 992.6 | 92.6 | 2178 | 9 | AF055008 | 9 | AF055008 Homo sapi |
| 6 | 991 | 92.4 | 2157 | 9 | AK000607 | 9 | AK000607 Homo sapi |
| 7 | 989.4 | 92.3 | 1630 | 9 | AK023348 | 9 | AK023348 Homo sapi |
| 8 | 747.2 | 69.7 | 1779 | 6 | ARI79655 | 6 | ARI79655 Sequence 3 |
| 9 | 747.2 | 69.7 | 1779 | 6 | II1830 | 6 | II1830 Sequence 3 |
| 10 | 629.2 | 58.7 | 139488 | 9 | AC003043 | 9 | AC003043 Homo sapi |
| 11 | 621.4 | 58.0 | 3944 | 9 | HUMGRANUL | 9 | HUMGRANUL |
| 12 | 596.4 | 55.6 | 180520 | 2 | AC019152 | 2 | AC019152 Homo sapi |
| 13 | 565 | 52.7 | 189391 | 2 | AC019151 | 2 | AC019151 Homo sapi |
| 14 | 560.6 | 52.3 | 2153 | 10 | RNEPIT1 | 10 | RNEPIT1 |
| 15 | 557.4 | 52.0 | 2113 | 10 | RATGABE12A | 10 | RATGABE12A |
| 16 | 546.4 | 51.0 | 2125 | 10 | MUSAG | 10 | MUSAG |
| 17 | 533.8 | 49.8 | 2137 | 6 | ARI75447 | 6 | ARI75447 Sequence |
| 18 | 532.2 | 49.6 | 2100 | 10 | MMEPIT1 | 10 | MMEPIT1 |
| 19 | 492 | 45.9 | 1867 | 9 | BC000324 | 9 | BC000324 Homo sapi |
| 20 | 469.2 | 43.8 | 1767 | 6 | ARI79654 | 6 | ARI79654 Sequence |
| 21 | 469.2 | 43.8 | 1767 | 6 | II1829 | 6 | II1829 Sequence 1 |
| 22 | 448.4 | 41.8 | 1767 | 6 | ARI79656 | 6 | ARI79656 Sequence |
| 23 | 448.4 | 41.8 | 1767 | 6 | II1831 | 6 | II1831 Sequence 5 |
| 24 | 364 | 34.0 | 55774 | 2 | AC091303 | 2 | AC091303 Homo sapi |
| 25 | 294.4 | 27.5 | 1952 | 10 | GPIAG | 10 | GPIAG |
| 26 | 267.8 | 25.0 | 8094 | 10 | MUSAP | 10 | MUSAP |
| 27 | 255.4 | 23.8 | 165490 | 2 | AC025326 | 2 | AC025326 Homo sapi |
| 28 | 255.4 | 23.8 | 220710 | 2 | AL596258 | 2 | AL596258 Mus muscu |
| 29 | 197.8 | 18.5 | 336 | 6 | AX198393 | 6 | AX198393 Sequence |
| 30 | 197.8 | 18.5 | 336 | 6 | AX208943 | 6 | AX208943 Sequence |
| 31 | 153.2 | 14.3 | 539 | 6 | ARI79657 | 6 | ARI79657 Sequence |
| 32 | 153.2 | 14.3 | 539 | 6 | II1832 | 6 | II1832 Sequence 7 |
| 33 | 111.4 | 10.4 | 113 | 11 | G54019 | 11 | G54019 922 Human H |
| 34 | 64.8 | 6.0 | 125020 | 9 | AF429315 | 9 | AF429315 Homo sapi |
| 35 | 64.6 | 6.0 | 341 | 6 | ARI79658 | 6 | ARI79658 Sequence |
| 36 | 64.6 | 6.0 | 341 | 6 | II1833 | 6 | II1833 Sequence 9 |
| 37 | 63.6 | 5.9 | 125020 | 9 | AF429315 | 9 | AF429315 Homo sapi |
| 38 | 61.2 | 5.7 | 18347 | 3 | CET22H2 | 3 | 281595 Caenorhabdi |
| 39 | 56.6 | 5.3 | 135369 | 2 | AC090557 | 2 | AC090557 Homo sapi |
| 40 | 56.6 | 5.3 | 173310 | 2 | AC090454 | 2 | AC090454 Homo sapi |
| 41 | 56.6 | 5.3 | 181923 | 9 | AC022872 | 9 | AC022872 Homo sapi |
| 42 | 44 | 4.1 | 110000 | 2 | LMFLCHR18_08 | 2 | Continuation (9 of |
| 43 | 43.4 | 4.0 | 131370 | 30 | AC017018 | 30 | AC017018 Homo sapi |
| 44 | 43.4 | 4.0 | 194236 | 2 | AL645944 | 2 | AL645944 Homo sapi |
| 45 | 43.2 | 4.0 | 202000 | 1 | AP000058 | 1 | AP000058 Aeropyrum |

ALIGNMENTS

| | | | | | | |
|------------|------------|-------------------------------------|---------|-----|--------|-----------------|
| RESULT 1 | ARI75456 | Sequence 16 from patent US 6309826. | 2095 bp | DNA | linear | PAT 17-DEC-2001 |
| LOCUS | ARI75456 | Sequence 16 from patent US 6309826. | | | | |
| DEFINITION | ARI75456 | Sequence 16 from patent US 6309826. | | | | |
| ACCESSION | ARI75456 | Sequence 16 from patent US 6309826. | | | | |
| VERSION | ARI75456.1 | GI:17916755 | | | | |
| KEYWORDS | Unknown. | | | | | |
| SOURCE | Unknown. | | | | | |
| ORGANISM | Unknown. | | | | | |

REFERENCE 1 (bases 1 to 2095)
AUTHORS Serrero, G.
TITLE 88kDa tumorigenic growth factor and antagonists
JOURNAL Patent: US 6309826-A 16 30-OCT-2001;
FEATURES
source Location/Qualifiers
1. 2095
/organism="unknown"
BASE COUNT 364 a 682 c 624 g 425 t
ORIGIN

Query Match 100.0%; Score 1072; DB 6; Length 2095;
Best Local Similarity 100.0%; Pred. No. 2.6e-234;

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/668,648
FILING DATE: 19910819
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212) 869-9741
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 1..537
US-07-668-648-7

Query Match 14.3%; Score 153.2; DB 1; Length 539;
Best Local Similarity 84.0%; Pred. No. 4.3e-29;
Matches 173; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 1 gggccccaccagtgccctggatggagagggccccagctcacctcagcctgcagacccca 60
DB 328 gggacccgccagtgccgtggatgaagaaagcccccacccacccctcagcctgcagaccc 387
QY 61 caagccttgaagagagatgccctgtgataatgtcagcagctgtccctccctccgatacc 120
DB 388 ggagcagtgaggggggagcgtccctgtgataagcgtcaccagctgtccttccactacc 447
QY 121 tgcgcacactcacgtctgggagtggggctgtgtcccaatcccagagggctgtctgtgc 180
DB 448 tgcgtgcgactcaagtcctggggagtgggcctgtgtcctccagagggctgtctgtgc 507
QY 181 tcggaccaccagcactgtgccccca 206
DB 508 tcggaccaccagcactgtgtcccca 533

Search completed: September 5, 2002, 18:22:29
Job time: 13352 sec


```
Matches 1072; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gggccccaccagtgctccctggatggagagggcccccagctcacctcagctgcccagaccaca 60
DB 1024 GGGCCCCACCAGTGCCCTGGATGGAGAGGGCCCCAGCTCACCTCAGCTGCCAGACCCA 1083
QY 61 caagccttgaagagagatgtccctctgtataatgtcagcagctgtccctccctccatacc 120
DB 1084 CAAGCCTTGAAGAGAGATGTCCTCTGTATATGTGACAGCTGTCCCTCTCCGATACC 1143
QY 121 tctcccaactcacctctggagagtgaggctgtctcaatcccaagagctgtctgctgc 180
DB 1144 TCTCCCAACTCACCTCTGGAGAGTGGGCTGTCTGCAATCCAGAGCTGTCTGCTGC 1203
QY 181 tcggaccaccagcactgtgccccagcgatcacagtgtgtagctgagggcagtgctcag 240
DB 1204 TCGGACCACCAGCAGTGTGCCCCCAGCATACAGTGTGTAGTGGGGCAGTGTCTAG 1263
QY 241 cgaggaagcagatcgtggtgactggaagatgctgcgcgcgcgcgcgcgcgcgcgcgc 300
DB 1264 CGAGGAAGCAGATGCTGCTGACTGGAGAGATGCTGCGCGCGCGCTTCTTATCC 1323
QY 301 cccccagagacatcgctgtaccagcacaccagctgcccgtggcggaacctgctgc 360
DB 1324 CACCCAGAGACATCGGCTGTACACAGCACACAGCTGCCCGGTGGCGGGAACCTGCTGC 1383
QY 361 ccgagccaggtgaggagctggcctgctcagctgccccatgctgtgtgctgcagagat 420
DB 1384 CCGAGCCAGGTGGAGAGCTGGGCTGCTGCCAGTGGCCCCATGCTGTGTGCTGCCAGGAT 1443
QY 421 cgcagcactgtgcccggctggctacacactgcaactggaagctcgatcctgcgagaag 480
DB 1444 CGCCAGCAGTGTGCCCGCTGGCTACACTGCACTGCAAGCTGATGCTGCGAGAG 1503
QY 481 gaagtgtctgtcccagcctgccaactctctgcccgttagcctcaactgaggtgtgaag 540
DB 1504 GAAGTGTCTGTGCCAGAGCTGCCACTTCTTGGCCCGGTAGCCCTCACCTGGGTGTGAAG 1563
QY 541 gacgtgagtgagggaagacactctccatgataaacacactgctgcccagagacaac 600
DB 1564 GACGTGGAGTGGGGAAGACACTTCTGCCATGATTAACAGACTGTCTGCCAGACAC 1623
QY 601 cgacagggctggcctgtctcctacgcccagggcgctgtgtgtgtgctgtagcgccac 660
DB 1624 CGACAGGCTGGCCCTGCTCCTACGCCAGGGCTGTGTGTGCTGTATCGGCCAC 1683
QY 661 tgcgtctcgtgtctcgtcgtgcagcagcaggggttaccagttgttgccagggagggc 720
DB 1684 TGTGTCTCTGTGCTGTCCGCTGCGCACGAGGGGTACCAAGTGTGTGCGCAGGGAGGC 1743
QY 721 ccgcgtggagcgcctcttgagggaccagccttgagcagctgtgtgagggacagta 780
DB 1744 CCGCGTGGGAGCCCTTTGAGGGACCCAGCTTGAGACAGCTGTGTGTGAGGGACAGTA 1803
QY 781 ctgaagactctgagcctcgggaccccaactcggaggggtgcccctgtcagcgcctccct 840
DB 1804 CTGAAGACTCTGAGCCCTCGGAGCCCACTCGAGGGGTGCCCTGTGCTCAGGGCTCCCT 1863
QY 841 agcacctcccccaataatctccctgagaccatctgagctcccatcacatggg 900
DB 1864 AGCACCTCCCCCAATAATCTCCCTGAGCCCAATCTTGAGCTCCCATCACCATGGG 1923
QY 901 agtggggctcaatcaatagggcctctccctgtcagagggggttgaggcaaaagccatt 960
DB 1924 AGTGGGGCTCAATCAATAGGCCCTTCCCTGTGTCAGAGGGGGTTGAGGCAAAAGCCATT 1983
QY 961 acaagtgccatccccctcccgtttcagtgaccctgtgtgcccaggtgtctttccctacc 1020
DB 1984 ACAAGTGGCATCCCCCTCCCGTTTCAGTGGACCCCTGTGCGCAGGTGTCTTCCCTATCC 2043
QY 1021 acagggatgtgtgtgaggtgtctttcaataaagtgttcactttctt 1072
DB 2044 ACAGGGGTGTGTGTGTGGGTGTGCTTCAATAAAGTTGTGCACTTCTT 2095
```

RESULT 2
HUMGRANUL
LOCUS
DEFINITION H.sapiens granulin mRNA, complete cds.
ACCESSION M75161
VERSION M75161.1 GI:183612
KEYWORDS
SOURCE Homo sapiens (tissue library: Clontech HL1058b) adult bone marrow mRNA.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 2095)
AUTHORS Bhandari, V., Palfree, R.G. and Bateman, A.
TITLE Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (5), 1715-1719 (1992)
MEDLINE 92179253
FEATURES
Location/Qualifiers
1..2095
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
/dev_stage="adult"
/tissue_lib="Clontech HL1058b"
1..2095
/gene="granulin"
13..1734
/gene="granulin"
/codon_start=1
/product="granulin"
/protein_id="AAA58617.1"
/db_xref="GI:183613"
/translation="MTLVSWALVAGVTRCPDGFVACCLDPGASVSCCRP
LDRWPTLSRLHGGPCQVDAHCSAGHSICFTVSGTSSCCPFPAVACGDGHHCPRG
FHCSDARSCTGSGNNNSVGAIQCDPSQDFCTCCVMDVSGCCPMPQASCCD
RVHCCPHGAFCDLVTRCTITPTGTHPLKLPKPAQTRNAVALSSVMCPDARSRCPD
STCCBLPSKYGCCPMPNATCCSHLHCCPQDVTCDLIQSKLSKNATDILLKLPA
HTVGVKCDMEVSCPDGTCRQASGAGCCPFTQAVCCEDHHCPCAGTCDTQKGT
CEQGHQVWHEKAPAHLSLDPDQALKRDVPCDNYSSCPSSDCCQLTSGEWGCCP
EAVCCSDHQHCPQRTVCVAGCQCGSEIVAGLEKMPARSGSLSHPRDGDQHTSC
PVGTCPCPQSGGSAWCCQLPHAVCCEDRHCPCAGTTCNVKARSKEVVSQAQATFL
ARSPHVGVKDECGEGHFCNDQTCRDNQGWACCPYAQGVCCADRRHCCPAGFRCA
RRGTLRLREAPNDALRDPALROLL"

sig_peptide 13..63
/gene="granulin"
/note="presumed signal peptide"
64..1791
mat_peptide
/gene="granulin"
/note="presumed mature peptide"
/product="granulin"
2075..2080
polya_signal
/gene="granulin"
BASE COUNT 364 a 682 c 624 g 425 t
ORIGIN

Query Match 100.0%; Score 1072; DB 9; Length 2095;
Best Local Similarity 100.0%; Pred. No. 2.6e-234; Mismatches 0; Indels 0; Gaps 0;
Matches 1072; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gggccccaccagtgctccctggatggagagggcccccagctcacctcagctgcccagaccaca 60
DB 1024 GGGCCCCACCAGTGCCCTGGATGGAGAGGGCCCCAGCTCACCTCAGCTGCCAGACCCA 1083
QY 61 caagccttgaagagagatgtccctctgtataatgtcagcagctgtccctccctccatacc 120
DB 1084 CAAGCCTTGAAGAGAGATGTCCTCTGTATATGTGACAGCTGTCCCTCTCCGATACC 1143
QY 121 tctcccaactcacctctggagagtgaggctgtctcaatcccaagagctgtctgctgc 180
DB 1144 TCTCCCAACTCACCTCTGGAGAGTGGGCTGTCTGCAATCCAGAGCTGTCTGCTGC 1203
QY 181 tcggaccaccagcactgtgccccagcgatcacagtgtgtagctgagggcagtgctcag 240
DB 1204 TCGGACCACCAGCAGTGTGCCCCCAGCATACAGTGTGTAGTGGGGCAGTGTCTAG 1263
QY 241 cgaggaagcagatcgtggtgactggaagatgctgcgcgcgcgcgcgcgcgcgcgcgc 300
DB 1264 CGAGGAAGCAGATGCTGCTGACTGGAGAGATGCTGCGCGCGCGCTTCTTATCC 1323
QY 301 cccccagagacatcgctgtaccagcacaccagctgcccgtggcggaacctgctgc 360
DB 1324 CACCCAGAGACATCGGCTGTACACAGCACACAGCTGCCCGGTGGCGGGAACCTGCTGC 1383
QY 361 ccgagccaggtgaggagctggcctgctcagctgccccatgctgtgtgctgcagagat 420
DB 1384 CCGAGCCAGGTGGAGAGCTGGGCTGCTGCCAGTGGCCCCATGCTGTGTGCTGCCAGGAT 1443
QY 421 cgcagcactgtgcccggctggctacacactgcaactggaagctcgatcctgcgagaag 480
DB 1444 CGCCAGCAGTGTGCCCGCTGGCTACACTGCACTGCAAGCTGATGCTGCGAGAG 1503
QY 481 gaagtgtctgtcccagcctgccaactctctgcccgttagcctcaactgaggtgtgaag 540
DB 1504 GAAGTGTCTGTGCCAGAGCTGCCACTTCTTGGCCCGGTAGCCCTCACCTGGGTGTGAAG 1563
QY 541 gacgtgagtgagggaagacactctccatgataaacacactgctgcccagagacaac 600
DB 1564 GACGTGGAGTGGGGAAGACACTTCTGCCATGATTAACAGACTGTCTGCCAGACAC 1623
QY 601 cgacagggctggcctgtctcctacgcccagggcgctgtgtgtgtgctgtagcgccac 660
DB 1624 CGACAGGCTGGCCCTGCTCCTACGCCAGGGCTGTGTGTGCTGTATCGGCCAC 1683
QY 661 tgcgtctcgtgtctcgtcgtgcagcagcaggggttaccagttgttgccagggagggc 720
DB 1684 TGTGTCTCTGTGCTGTCCGCTGCGCACGAGGGGTACCAAGTGTGTGCGCAGGGAGGC 1743
QY 721 ccgcgtggagcgcctcttgagggaccagccttgagcagctgtgtgagggacagta 780
DB 1744 CCGCGTGGGAGCCCTTTGAGGGACCCAGCTTGAGACAGCTGTGTGTGAGGGACAGTA 1803
QY 781 ctgaagactctgagcctcgggaccccaactcggaggggtgcccctgtcagcgcctccct 840
DB 1804 CTGAAGACTCTGAGCCCTCGGAGCCCACTCGAGGGGTGCCCTGTGCTCAGGGCTCCCT 1863
QY 841 agcacctcccccaataatctccctgagaccatctgagctcccatcacatggg 900
DB 1864 AGCACCTCCCCCAATAATCTCCCTGAGCCCAATCTTGAGCTCCCATCACCATGGG 1923
QY 901 agtggggctcaatcaatagggcctctccctgtcagagggggttgaggcaaaagccatt 960
DB 1924 AGTGGGGCTCAATCAATAGGCCCTTCCCTGTGTCAGAGGGGGTTGAGGCAAAAGCCATT 1983
QY 961 acaagtgccatccccctcccgtttcagtgaccctgtgtgcccaggtgtctttccctacc 1020
DB 1984 ACAAGTGGCATCCCCCTCCCGTTTCAGTGGACCCCTGTGCGCAGGTGTCTTCCCTATCC 2043
QY 1021 acagggatgtgtgtgaggtgtctttcaataaagtgttcactttctt 1072
DB 2044 ACAGGGGTGTGTGTGTGGGTGTGCTTCAATAAAGTTGTGCACTTCTT 2095


```
QY 1 gggccccacagtgccctggatggagagggccccagctcaactgaagcctgcccagaccac 60
Db 1052 GGGCCCCACAGTGCCTGGATGGAGAGGGCCCCAGCTCACCTCAGCTGCCAGACCCA 1111
QY 61 caagccttaagagagatccccctgtataatgtcagcagctgtccctccctccagatcc 120
Db 1112 CAAGCCTTGAAGAGAGATGCCCTGTGATATGTGATGATGATGATGATGATGATGAT 1171
QY 121 tgcgtcaactcagctgtgggagtggtggtgctgtgtccaatccagagagctgtctgtgc 180
Db 1172 TGCTGCCAATCAGCTGTGGGAGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1231
QY 181 tcggaccacagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
Db 1232 TCGACACACAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1291
QY 241 cggaggaagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
Db 1292 CGAGGAAGCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1351
QY 301 caccagagacatcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
Db 1352 CACCCAGACATCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1411
QY 361 ccgagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
Db 1412 CCGAGCCTGGTGGGAGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1471
QY 421 cgcagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480
Db 1472 CGCCAGACACTGCTGCCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1531
QY 481 gaagtgtctctccagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
Db 1532 GAAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1591
QY 541 gacgtgagtggtgggagagacacttctgcatgataaacagacactgtgtgtgtgtgtgt 600
Db 1592 GACGTGAGTGTGGGAGAGACACTTCTGCATGATATAACAGACCTGTGTGTGTGTGT 1651
QY 601 cgcagggctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
Db 1652 CGACAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1711
QY 661 tgcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
Db 1712 TGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1771
QY 721 ccgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
Db 1772 CCGGCTGGGAGCGCCCTTTGAGGAGCCAGCCTTGAGACAGCTGTGTGTGTGTGTGTGT 1831
QY 781 ctgaagactgtgaagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
Db 1832 CTGAAGACTGTGACGCTGTGGGAGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1891
QY 841 agcactcccccaacaaattctccctgtgacccattctgagctcccccaacatcaggg 900
Db 1892 AGCAGCTTCCCTTAACCAATCTCCCTGTGACCCCATCTTGTGAGCTTCCCATCAGG 1951
QY 901 aggtggggtcctcaatcagggccttccctgtcagaaggggtgtgtgtgtgtgtgtgtgt 959
Db 1952 AGGTGGGCTCAATCTAAGG-CCTTCCCTGTGAGAGGGGGTGTGTGTGTGTGTGTGTGT 2010
QY 960 tacaagtgtccatccccctccccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1019
Db 2011 TACAAGTGTCCATCCCTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2070
QY 1020 cacaggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1038
Db 2071 CACAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2089
```

```
RESULT 5
AF055008
LOCUS Homo sapiens clone 24720 epithelin 1 and 2 mRNA linear PRI 02-APR-1998
DEFINITION Homo sapiens clone 24720 epithelin 1 and 2 mRNA, complete cds.
ACCESSION AF055008
VERSION AF055008.1 GI:3005729
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2178)
AUTHORS Andersson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
REFERENCE 2 (bases 1 to 2178)
AUTHORS Yu,W., Andersson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
Ricafrente,J.Y., Wentland,M.A., Lennop,G. and Gibbs,R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 97264341
REFERENCE 3 (bases 1 to 2178)
AUTHORS Yu,W. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
Location/Qualifiers
1..2178
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE Consortium 24720"
/sex="female"
/tissue_type="brain"
/clone_lib="1N1B"
/dev_stage="infant"
63..1844
/notes="similar to Homo sapiens epithelin 1 and 2 encoded
by GenBank Accession Number X62320"
/codon_start=1
/product="epithelin 1 and 2"
/protein_id="AAC09359.1"
/db_xref="GI:3005730"
/translation="MWTLVSWALTAGLVAGTRCPDQGFQCPVACCLDPGASYSCCRPR
LIDKWPITLRLHGLGCPQVDAHCSAGHSCLFTVSGSCFFPEAVACGHHCCPRG
FHCADGRSCFQRGNSVGAIQCPDQFECDFSTCCVMVDGSGCCPMPQASCCED
RVHCCPHGAFCDLVHTRCTITPTHTPLAKLPARTNRAVALSSVMCPDARSRCPDG
STCCCLPSKYGCCPNATCCSDHLHCCPQDTCVCLIQSKLSKENATDDLTKLPA
HTVGDVKDNEVSPDGYCCRLQSGANGCCPFTQAVCCEDHILHCCPAGTCPTQKGT
CBGPHQVPMMEKAPAHLSLPDPAKLRDVPDNCVSCPSDSTCCCLTSGEWCCPILP
EAVCCSDHQHCCPQGYTCVAGQCQRGSEIVAGLEKMPARRASLSHPRDITGDOHTSC
PVGOTCCPSLIGGSWACQLPHAYCCEDRQHCPCAGTCTCNKARSCKEVYSAQPAFL
ARSPHGVKDVCEGEGHFDNCTCCRDNRQGWACCPYRGVCCARRHRCPCAFRCA
ARSTKCLRREAPNDAPLRDPAIRLL"
BASE COUNT 401 a 637 c 646 g 434 t
ORIGIN
Query Match 92.6%; Score 992.6; DB 9; Length 2178;
Best Local Similarity 98.5%; Pred. No. 3.3e-216;
Matches 1023; Conservative 0; Mismatches 14; Indels 2; Gaps 2;
QY 1 gggccccacagtgccctggatggagagggccccagctcaactgaagcctgcccagaccac 60
Db 1074 GGGCCCCACAGTGCCTGGATGGAGAGGGCCCCAGCTCACCTCAGCTGCCAGACCCA 1133
QY 61 caagccttaagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
Db 1134 CAAGCCTTGAAGAGAGATGTCCCTGTGATATGTGATGATGATGATGATGATGATGAT 1193
QY 121 tgcgtcaactcagctgtgggagtggtggtgctgtgtccaatccagagagctgtctgtgc 180
Db 1194 TGCTGCCAATCAGCTGTGGGAGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1253
```

```
QY 181 tcggaccaccagcactgctgccccagcagatcacagctgtgtagctgagggcagtgctag 240
Db 1254 TCGGACCACAGCAGCTGCTGCCCGAGGCTACAGCTGTGTAGCTGAGGGCAGGTGTGAG 1313
QY 241 cagagaaacgagatgctgctgagctgagagagatgctgccccgcgctgcttctatcc 300
Db 1314 CGAGGAACCGAGATGCTGCTGGACTGGAGATGAGAGATGCTGCCCGCGGCTTCCTTATCC 1373
QY 301 caccacagagacatcgctgtgaccagcacaccagctcccggtgagcggaacctgctgc 360
Db 1374 CACCCAGAGACATCGCTGTGTGACGACACACAGCTGCCCGGTGGCGCAGACCTGCTGC 1433
QY 361 ccgagccaggggtggagctggcctgctgcccagttgccccatgctgtgtgtgagagat 420
Db 1434 CCGAGCCCTGGGTGGAGCTGGCCCTGCTGCCAGTTGGCCCCATGCTGTGTGTGGAGGAT 1493
QY 421 cgcagcactgctgccccgctggtgtacacacctgcaacgtgaaagctcgatctcgcagaga 480
Db 1494 CGCCAGCAGCTGCTGCCCGGCTGGCTTACACCTGCAACGTTGAAGGCTCGATCTTCGCGA 1553
QY 481 aaagtgtctctgcccagctgcccactctctgcccagctgcccagctgcccagctgccc 540
Db 1554 GAAGTGGTCTCTGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1613
QY 541 gacgtggagtggtggaaaggacactctgcatgataaccagacactgctgcccagagaca 600
Db 1614 GACGTGGAGTGTGGGAGGAGACATCTTCCCATGATATACAGACCTGCTGCCGAGACAA 1673
QY 601 cgaagggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 660
Db 1674 CGACAGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1733
QY 661 tctctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 720
Db 1734 TGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1793
QY 721 ccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 780
Db 1794 CCGCGCTGGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1853
QY 781 ctgagactctgagcctctgagcctctgagcctctgagcctctgagcctctgagcctct 840
Db 1854 CTGAAGACTCTGAGCCTCTGAGCCTCTGAGCCTCTGAGCCTCTGAGCCTCTGAGCCT 1913
QY 841 agcactcccttaacaaattctccctgagcctctgagcctctgagcctctgagcctct 900
Db 1914 AGCAGCTCCCTTAACCAAAATCTCCCTGGAGCCCAATCTGAGCTCCCAATCAGCATGG 1973
QY 901 aggtgggctcaatctaaagcctctccctgagcctctgagcctctgagcctctgagcct 959
Db 1974 AGGTGGGCTCAATCTAAGG-CTTCCCTGTGTCAGAGGGGCTTGGCAAGGCCACAT 2032
QY 960 tacaagctgcaatccctctccctctgagcctctgagcctctgagcctctgagcctct 1019
Db 2033 TACAAGCTGCCATCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2092
QY 1020 cacaggggtgtgtgtgt 1038
Db 2093 CACAGGGGTGTGTGTGTGT 2111

RESULT 6
AC000607 AK000607 2157 bp mRNA linear PRI 22-FEB-2000
LOCUS Homo sapiens cDNA FLJ20600 fis, clone KAT07521, highly similar to
DEFINITION AF055008 Homo sapiens clone 24720 epithelin 1 and 2 mRNA.
AC000607 AK000607
VERSION AK000607.1 GI:7020815
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to
mRNA, clone_lib:KAT clone:KAT07521.
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
2 (bases 1 to 2157)
REFERENCE
AUTHORS Unpublished (2000)
TITLE NEDO human cDNA sequencing project
JOURNAL
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
JOURNAL Direct Submission
Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo, Deptment
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass-sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source Location/Qualifiers
1..2157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone="KAT07521"
/clone_lib="KAT"
/note="cloning vector pME18SFL3"
misc_feature 1..2157
/note="highly similar to AF055008 Homo sapiens clone 24720
epithelin 1 and 2 mRNA"
BASE COUNT 399 a 691 c 639 g 428 t
ORIGIN

Query Match 92.4%; Score 991; DB 9; Length 2157;
Best Local Similarity 98.4%; Pred. No. 7.7e-216;
Matches 1022; Conservative 0; Mismatches 15; Indels 2; Gaps 2;

QY 1 gggccccaccagtgctgctggatggagagggccagctcacctcagctgcccagacc 60
Db 1052 GGGCCCCACCAGTGGCTGGATGGAGAGAGGCCAGCTCAGCTCAGCTGCCAGACCA 1111
QY 61 caagccttgaagagagatgtccctgtgataatgtcagcagctgtccctcccgatacc 120
Db 1112 CAAGCCTTGAAGAGAGATGTCCCTGTGATAATGTACAGCAGCTGTCCCTCCGATACC 1171
QY 121 tctgtccaaactcaactctggggagtggggctgtctgtccaaatccagagagctgtct 180
Db 1172 TGCTGCCAACTACACTCTGGGGAGTGGGGCTGTCTGTCCAAATCCAGAGAGCTGTCTGCTGC 1231
QY 181 tcggaccaccagcactgctgccccagcagatcacagctgtgtagctgagggcgagtgct 240
Db 1232 TCGGACCACCAGCAGCTGCTGCCCGAGGCTACAGCTGTGTAGCTGAGGGGAGTGTGAG 1291
QY 241 cagagaaacgagatcgtgctgagctggagagagatgctgccccgcgctgcttctatcc 300
Db 1292 CGAGGAACCGAGATGCTGCTGGACTGGAGAGATGCTGCCCGCGGCTTCCTTATCC 1351
QY 301 caccacagagacatcgctgtgaccagcacaccagctcccggtgagcggaacctgctgc 360
Db 1352 CACCCAGAGACATCGCTGTGTGACGACACACAGCTGCCCGGTGGCGCAGACCTGCTGC 1411
QY 361 ccgagccaggggtggagctggcctgctgcccagttgccccatgctgtgtgtgagagat 420
Db 1412 CCGAGCCCTGGGTGGAGCTGGCCCTGCTGCCAGTTGGCCCATGCTGTGTGTGCGAGGAT 1471
QY 421 cgcagcactgctgccccgctggtgtacacctgcaacgtgaaagctcgatctcgcagaga 480
```

```

|||||
Db 1472 CGCCAGCACTGCTGCCGGCTGGCTACACCTGCAACGTGAAGGCTCGATCTCGAGAA 1531
Qy 481 gaagtgtcttgcagcctgcacacttctcctgctcctgagcctcagctggtgtaag 540
Db 1532 GAAGTGTCTCTGCCAGCCTGCCACCTTCCTGCGCGTAGCCCTCACGTGGGTGAAG 1591
Qy 541 gactgtgagtggtgggaagacacttctgcatgataaacagacactgtgctgcagacaac 600
Db 1592 GACCTGGAGTGTGGGAGGACACTTCTGCATGATACACGACCTGCTGCCGAGACAAC 1651
Qy 601 cgacagggctgggctgtctcctcagcctcaggggctgtgtgtgtgtgtgtgtgtgtgt 660
Db 1652 CGACAGGCTGGGCTGTCTGCTCCCTACCGCAGGGGCTCTCTTGTGTGTGTGTGTGT 1711
Qy 661 tgetctcgtgtgtcctcctgctgcagcaggggtaccagtgttgcagagggcc 720
Db 1712 TGCTGTCTGCTGCTTCCGCTGCGCAGCCAGGGGTACCAAGTGTGTTCGCGAGGAGGCC 1771
Qy 721 ccgcgtgggagcgccttttggagaccagccttgagacagctgctgtgaggacagta 780
Db 1772 CCGCGCTGGGAGCCCTTTGAGGAGCCAGCCCTTGAGACAGCTGCTGTGAGGACAGTA 1831
Qy 781 ctgaagactgtcagcctcctggagccacacacacacacacacacacacacacacac 840
Db 1832 CTGAAGACTCTGCAGCCTCGGAGCCACCTCGGAGGGTGCCTCTGCTCAGGCTCCT 1891
Qy 841 agcactcctcctaaccaattctcctgagccacacacacacacacacacacacacac 900
Db 1892 AGCACTCTCCTTAACCAAAATCTCCCTGAGCCACCAATCTGAGCTCCCATCACCATGG 1951
Qy 901 aggtgggctcctaatcgaagcctcctcctcctcctcctcctcctcctcctcctcct 959
Db 1952 AGTGGGCTCTCACTAAGG-CTTCTCTCTCAGAGAGGGGTTGTGCAAAAGCCACAT 2010
Qy 960 tacaagctgcacacacacacacacacacacacacacacacacacacacacacacacac 1019
Db 2011 TACAAGCTGCACCTCCCTCCCGCTTCAAGTGAGACCTGTGGCCAGGCTGCTTTTCCCTATC 2070
Qy 1020 cacaggggtgtgtgtgt 1038
Db 2071 CACAGGGGTGTGTGTGT 2089

RESULT 7
AK023348
LOCUS Homo sapiens CDNA FLJ13286 fis, clone OVARC1001154, highly similar
DEFINITION to Homo sapiens clone 24720 epithelin 1 and 2 mRNA.
ACCESSION AK023348
VERSION AK023348.1 GI:10435243
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens ovary, tumor tissue CDNA to mRNA, clone_lib:OVARC1
clone:OVARC1001154.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1630)
AUTHORS Isoqai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sato,H., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,T., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1630)
AUTHORS Isoqai,T. and Otsuki,T.
JOURNAL Direct Submission
TITLE Submitted (23-AUG-2000) Takao Isoqai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)

```

```

COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source 1. 1630
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1001154"
/tissue_type="ovary, tumor tissue"
/clone_lib="OVARC1"
/note="cloning vector: pME18SFL3"
85. .1326
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="BABI4535.1"
/db_xref="GI:10435244"
/translaton="MAITAAAGASTAVQTGDPASKQVTPWPSSALIVSSNARTSP
KYLWMSAPGGAAPCPRLPAVKTGCTAVCDLIQSKLCKENATDLTKLPAHTVGDV
KDMVECPDGYTCRLOSAGWGCPPTQAVCCEDHHCOPAGFTCTQKGTCTGQGGPH
QVPMERKAPAHLSLPDQALKRDVPCDNVSCPSDSTCCOLTSCEWCCPTPEAVCCS
DHQCCPGTCTVABGQCQRGSEIVAGLEKMPARRASLSRPRDJGCCQHPSCPVGTC
CPSLGGSWACQULPHAVCCEDRQHCCPAGYTCNVKARSEKEVVSQAFTFLARSPHV
GVKDECGEGHFDHNDTCRDNRQGWACCPYRGVCCADRRHRCPPAGFAARGTKC
LRREAPRWDAPLRDPALRQLL"
misc_feature 297..298
/note="260 bases segment is present in AF055008, X62320
and AK000607."
BASE COUNT 285 a 517 c 489 g 339 t
ORIGIN
Query Match 92.3%; Score 989.4; DB 9; Length 1630;
Best Local Similarity 98.3%; Pred. No. 1.8e-215;
Matches 1021; Conservative 0; Mismatches 16; Indels 2; Gaps 2;
Qy 1 gggccccaccagtgctcctggatggagagggccccagctcactcagctgcagacc 60
Db 556 GGGCCCCACCAGGTGCCCTGGATGGAGAGGCCCCAGCTCACTCAGCCTGCAGACCCA 615
Qy 61 caagccttgaagagagatgtccctgtgataatgtcagcagctgtccctcccgatacc 120
Db 616 CAAGCCTTGAAGAGAGATGTCCCTGTGATAATGTACAGAGCTGTCCCTCCGATACC 675
Qy 121 tgcctccaactcagctctggagtgaggctgctctccaatcccagaggtgctgtgctgc 180
Db 676 TGCTGCCAACTCACGTCTGGGGAGTGGGGCTGCTGTCCAAATCCAGAGGCTGTCTGCTGC 735
Qy 181 tcggaccaccagcactgtgtccccccagcagatatacacgtgtgtagctgagggcagtgacg 240
Db 736 TCGGACCACCAGCACTGTGCTGCCCGCCAGGGCTACACGTGTGTAGCTGAGGGCAGTGTGAG 795
Qy 241 cgaggaagcagagatcgtggtgactggagagagagctgtgcgcgcgcgcgcgcgcgcgcgc 300
Db 796 CGAGGAAGCGAGATCGTGGCTGGAGTGGAGAGATGCTGTGCCCGCGGGCTTCCTTATCC 855
Qy 301 caccacagagacatcgctgtgtaccagacacacacacacacacacacacacacacacacac 360
Db 856 CACCCAGAGACATCGGCTGTGACACAGACACACAGCTGCCCGTGGGGGAGAGACCTGCTGCTGC 915
Qy 361 ccgagccaggggtggagctggcctgctgctcagctgtccacatgtgtgtgtgtgtgtgtgtgt 420
Db 916 CCGAGCCTGGGTGGAGCTGGGCTGCTGCCAGTTGCCCATGCTGTGTGTGTGTGTGTGTGTGT 975
Qy 421 cgccagcactgtgtccccgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480
Db 976 CGCCAGCACTGTGTGCCGGCTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1035
Qy 481 gaagtgtgtctgtgccagcagcctgccacacacacacacacacacacacacacacacacacac 540

```

| | | | |
|----|------|---|------|
| Db | 1132 | TGCTGCCAACTCAGCTGSGGAGTGGGCTGCTGTGCCAATCCAGAGGCTGTCTGTGCTGC | 1191 |
| Qy | 181 | tcggaccaccagcaactgctgccccccagcgtatacaacgtgtgtagtgcgagggcagtgctcag | 240 |
| Db | 1192 | TCGGACCACCAGACACTGCTGCCCCAGGCTTACACGTGTGTAGCTGTAGCTGTAGGGCAGTGTCTAG | 1251 |
| Qy | 241 | cgaggaaqcgagatcctggctggactggagaagatgcctgcccccgcggtttctctatccc | 300 |
| Db | 1252 | CGAGGAAGCAGATCGTGGCTGGACTGGAGAAGATGCTGCCCGCGGGCTTCCTTATATCC | 1311 |
| Qy | 301 | caccaccagagacatcggtctgtaccagcacaccagctgcccggtggcggaacctgctgc | 360 |
| Db | 1312 | CACCCACAGACATCGGCTGTGACACAGCACACCACTGCTGCCGGTGGGCGACACCTGCTGTC | 1371 |
| Qy | 361 | ccgagccagggtggagctgggctgctgccagttgcccatctgtgtgtctgctgcgaggaat | 420 |
| Db | 1372 | CCGAGCCTGGGTGGAGCTGGGCTGCTGCCAGTTGCCCCATGCTGTGTGCTGGCAGGAT | 1431 |
| Qy | 421 | cgccagcaactctgcgcggctggctacacctgcaacgtgaaggctcgatctctgcgagaag | 480 |
| Db | 1432 | CGCCAGCACTGCTGCCCGGTGGCTACACTGCACGTGAAGGCTCGATCCTCTGCCAGAA | 1491 |
| Qy | 481 | gaagtgtctctgccacagcctgccacctctctggcccgtagccctcaacgtgggtgtgaag | 540 |
| Db | 1492 | GAAGTGTGCTCTGCCAGCCTGCACCTTTCCTGGCCCGTAGCCCTCAGCTGGGTGTGAAG | 1551 |
| Qy | 541 | gaagtgaagtgggaaggacacttctgccatgataaccagacactgctgcgagacaac | 600 |
| Db | 1552 | GACGTGAGTGTGGGAAGGACACTTCTGCCATGATTAACACACTGCTGCCGAGACAA | 1611 |
| Qy | 601 | cgacaggctggggcctgctgtccctacgcccaggcgctgtgtgtgtgatcgcgccac | 660 |
| Db | 1612 | CGACAGGGCTGGGCTGCTGTCCCTACCGCCAGGGGCTGTGTGTGCTGATCGGCGCCAC | 1671 |
| Qy | 661 | tgtctcctgctggtcttcogctgcgcacgcagggtgtaccaagtgtttgcgagggagggcc | 720 |
| Db | 1672 | TGCTGTCTGCTGCTTCGCTGCGCAGCCAGGGTACCAAGTGTTCGACAGGAGGCC | 1731 |

| | | | |
|-----------------------|--|---|----------------------------|
| Qy | 721 | ccgcgtgggagcgccttttgaggaccacgccttgagacagctgctg | 768 |
| | | | |
| Db | 1732 | CGCGCTGGGACGCCCTTTGAGGGACCCAGCCTTGACACAGCTGCTG | 1779 |
| | | | |
| RESULT | 9 | | |
| LOCUS | II1830 | 1779 bp | DNA linear PAT 26-JUL-1995 |
| DEFINITION | Sequence 3 from Patent US 5416192. | | |
| ACCESSION | II1830 | | |
| VERSION | II1830.1 | GI:909273 | |
| KEYWORDS | Unknown. | | |
| SOURCE | Unknown. | | |
| ORGANISM | Unclassified. | | |
| REFERENCE | 1 (bases 1 to 1779) | | |
| AUTHORS | Shoyab M. and Plowman G.D. | | |
| TITLE | Epithelins: novel cysteine-rich growth modulating proteins | | |
| JOURNAL | Patent: US 5416192-A 3 16-MAY-1995; | | |
| FEATURES | Location/Qualifiers | | |
| source | 1. .1779 | | |
| | /organism="unknown" | | |
| BASE COUNT | 304 a 578 c 547 g 350 t | | |
| ORIGIN | | | |
| Query Match | 69.7%; | Score 747.2; | DB 6; |
| Best Local Similarity | 98.3%; | Pred. No. 3.2e-160; | |
| Matches | 755; | Conservative 0; | Mismatches 13; |
| | | Indels 0; | Gaps 0; |
| Qy | 1 | ggggccaccaggtgccttgatggagaagggcccccagctcacctcagcttcagcctgacgaccca | 60 |
| | | | |
| Db | 1012 | GGGGCCCCACCAGGTGCCCTGGATGGGAAGGCCCCAGCTACCTCAGCTGCCAGACCCA | 1071 |
| | | | |

| | | | |
|----|------|--|------|
| Qy | 61 | caagccttgaagagagatgtccctctgtgataatgtcgaagcagtgtccctctccgtacatacc | 120 |
| Db | 1072 | CAAGCCTTGAAGAGAGATGTCCCTCTGTGATATGTATGTCAAGACGTCTCCCTCTCTCCGATAACC | 1131 |
| Qy | 121 | tgtctgccaactcactctgaggagtggggctgctgtccaatccacagaggtgtctctgtcgc | 180 |
| Db | 1132 | TGCTGCCACTCACGTCTGGGAGGTGGGGCTGTCTCATATCCACAGAGGCTGTCTGCTGC | 1191 |
| Qy | 181 | tcgggaccacagcactgtgtcccccagcgatacacagtgtgtagctgagggggcagtgctcag | 240 |
| Db | 1192 | TCGGACACACAGCACTGCTGCCCCCCAGGGCTACAGCTGTGTAGCTGAGGGCGAGTGTCAAG | 1251 |
| Qy | 241 | cggagaagcgaatcgtgactgtgactcgagaagaatgcctgcccccgcgattcctatcc | 300 |
| Db | 1252 | CGAGGAAGCGAGATCTGGCTGGATGGAGAAGATGCTGCCGCCGGGCTTCTTTATCC | 1311 |
| Qy | 301 | caccccagagacatcggctgtgaccagcacaccagctgccgggtgggcgaacctgctgc | 360 |
| Db | 1312 | CACCCAGAGACATCGGCTGTACACAGCACACAGCTGCCCGTGGGCGAGACCTGTCTGC | 1371 |
| Qy | 361 | ccgagccaggggtggagctgggctgtgcgaattgcgaattgcacctgtgtgtctgcgaggat | 420 |
| Db | 1372 | CCGAGCCTGGGTGGGAGCTGGGCCCTGCTGCCAGTTGCCCATGTGTGTGCTGCGAGGAT | 1431 |
| Qy | 421 | cgcagcactgtgccggctggctcacactgcaacgtgaaggctcgatcctctgcgagaag | 480 |
| Db | 1432 | CGCCAGCACTGCTCCCGGCTGGCTACACTGCACGTGAAGGCTCGATCTCTGCGAGAAG | 1491 |
| Qy | 481 | gaagtggctctgccagcctgccacctctctggcccctagacctcaogtgggtgtgtaag | 540 |
| Db | 1492 | GAAGTGTGCTCTGCCAGCGCTGCCACTTCTCTGGGCCCTAGCCCTACGCTGGGTGTGAAG | 1551 |
| Qy | 541 | gaagtgaagtgggaagagacactctgcacatgaaccagacctgtcgcgagacaac | 600 |
| Db | 1552 | GAGCTGGAGTGTGGGAAGGACACTTCTGCCATGATACACAGACCTGCTGCCGAGACAAC | 1611 |
| Qy | 601 | cgacagggtgggacctgtctccctacgcccaggcgctgtgtgtgtgtatcgcgccac | 660 |
| Db | 1612 | CGACAGGGCTTGGGCCCTGCTGTCCCTACCGCCAGGGCGCTGTGTGTGTGTATCTGGCGGCAC | 1671 |
| Qy | 661 | tgtgtcctctgagcttcgctgcgcagcagggtaccaaagtgtttgcgcaggagccc | 720 |
| Db | 1672 | TGCTGTCTCTGGCTTCCGCTGCGCAGCCAGGGGTACCAAGTGTTTGCGCAGGGAGGCC | 1731 |
| Qy | 721 | ccgcgctgggagccccttttgaggaccacagcctttgacacagctgctg | 768 |
| Db | 1732 | CCGCGCTGGAGACCCCTTTTGGAGGACCCAGCCTTTGAGACAGCTGCTGTG | 1779 |

| | |
|---------|--------------|
| exon | 1072. .1184 |
| | /number-4 |
| intron | 1185. .1284 |
| | /number-4 |
| exon | 1285. .1421 |
| | /number-5 |
| intron | 1422. .1533 |
| | /number-5 |
| exon | 1534. .1641 |
| | /number-6 |
| intron | 1642. .1875 |
| | /number-6 |
| exon | 1876. .2000 |
| | /number-7 |
| intron | 2001. .2199 |
| | /number-7 |
| exon | 2200. .2297 |
| | /number-8 |
| intron | 2298. .2376 |
| | /number-8 |
| exon | 2377. .2621 |
| | /number-9 |
| intron | 2622. .2840 |
| | /number-9 |
| exon | 2841. .3074 |
| | /number-10 |
| intron | 3075. .3182 |
| | /number-10 |
| exon | 3183. .3413 |
| | /number-11 |
| intron | 3414. .3498 |
| | /number-11 |
| exon | 3499. .3944 |
| | /number-12 |
| ! COUNT | 691 a 1186 c |
| IN | |

| | | | | | | | |
|-----------------------|--|--------------|-----------|------------|-----|--------|-------|
| Query Match | 58.0%; | Score | 621.4; | DB | 9; | Length | 3944; |
| Best Local Similarity | 80.4%; | Pred. No. | 1.4e-131; | | | | |
| Matches | 866; | Conservative | 0; | Mismatches | 16; | Indels | 195; |
| Gaps | | | | | | | |
| 157 | ccaatccagaggctgctgctcgagaccacacagcactgctgcccacagcgatacag | 216 | | | | | |
| 2829 | CCACCCCCAGGCTGTGCTGTGCGACACACAGCACTGCTGCCCCAGGCTACAG | 2888 | | | | | |
| 217 | tgtgtagctgaggggcagtgctcagcgaggaagcagatcgtggtcgactggsagaagt | 276 | | | | | |
| 2889 | TGTGTAGCTGAGGGGCAGTGTGACGGAGAAAGAGATCGTGGTGTGACTGGAGAAGATG | 2948 | | | | | |
| 277 | ctgtcccgccgcggttcottatccaccccaagacacatcggtctgtaccagcacaccag | 336 | | | | | |
| 2949 | CTGTCCCGCGGGCTTCCTTATCCACCCACAGACATCGGGCTGTACCAACACCCAGC | 3008 | | | | | |
| 337 | tgcgcggttggcgcgaacctgctgcccgacaggtgtggagctgtggtcgtgccagttg | 396 | | | | | |
| 3009 | TGCCCCGTGGGCACACCTGCTGCCCGAGCTGGGTGGAGCTGGGCTTCTGCCAGTTG | 3068 | | | | | |
| 397 | ccccat----- | 402 | | | | | |
| 3069 | CCCCATGTAGTGCCTCCTCGCTGCCCTGTGATAGGGGAGCTAAGCCCACTGAGGGAGC | 3128 | | | | | |
| 403 | ----- | 408 | | | | | |
| 3129 | TAAGCCACGTGANGACAGGAACATATATGCCATTCTGTGTCCCTTCCGCCAGGCTGTG | 3188 | | | | | |
| 409 | tgtctcgaggatcgccagcactgctgcccggtgctgctacacctgcaactggaaggtcga | 468 | | | | | |
| 3189 | TGCTGGAGGATCGCCAGCACTGCTGCCGGTGGCTACACTTCAAGCTGAAGGCTCGA | 3248 | | | | | |
| 469 | tctcgagaagaagtgtctctgtcccagcgtgccaccttctgtgccgtgacctctcac | 528 | | | | | |
| 3249 | TCTTCGAGAAAGAGTGTCTCTGCCAGCGTGCACCTTCTTGGCCGTGACCCCTCAC | 3308 | | | | | |

| | | | | |
|---|----|---------|-------------------|---------------|
| * | i. | 1099: | contig of 1099 | bp in length |
| * | | 1199: | gap of unknown | length |
| * | | 1200: | contig of 1269 | bp in length |
| * | | 2468: | contig of unknown | length |
| * | | 2469: | gap of unknown | length |
| * | | 2569: | contig of 1579 | bp in length |
| * | | 4148: | gap of unknown | length |
| * | | 4248: | gap of unknown | length |
| * | | 4339: | contig of 1092 | bp in length |
| * | | 5340: | gap of unknown | length |
| * | | 5439: | contig of 1058 | bp in length |
| * | | 6498: | gap of unknown | length |
| * | | 6598: | contig of 1577 | bp in length |
| * | | 8175: | gap of unknown | length |
| * | | 8275: | contig of 6297 | bp in length |
| * | | 14572: | gap of unknown | length |
| * | | 14672: | gap of unknown | length |
| * | | 24767: | contig of 10095 | bp in length |
| * | | 24867: | gap of unknown | length |
| * | | 39139: | contig of 14273 | bp in length |
| * | | 39239: | gap of unknown | length |
| * | | 39240: | contig of 23562 | bp in length |
| * | | 62802: | gap of unknown | length |
| * | | 113454: | contig of 50553 | bp in length |
| * | | 113455: | gap of unknown | length |
| * | | 180520: | contig of 66966 | bp in length. |

[illegible]

```
1. 1099
   /note/assembly_name:Contig22
1200. 2468
   /note/assembly_name:Contig27*
2569. 4147
   /note/assembly_name:Contig31*
4248. 5339
   /note/assembly_name:Contig32*
5440. 6497
   /note/assembly_name:Contig33*
6598. 8174
   /note/assembly_name:Contig34*
8275. 14571
   /note/assembly_name:Contig35*
14672. 24766
   /note/assembly_name:Contig36
clone_end:SP6
vector_side:right*
24867. 39139
   /note/assembly_name:Contig37*
39240. 62801
   /note/assembly_name:Contig38*
62902. 113454
```

[illegible]

[illegible]

| | | | |
|----|------|---|------|
| Db | 1791 | AACGACTAAGAAACTCCACAGTCTCTGGGAACCCCTGTTCTTGAGGGTATCCACCACTCAGGC | 1850 |
| QY | 835 | ctccctagcacctccccctcaacaaattctccctggaccctattctgag-ctccccatca | 893 |
| Db | 1851 | CTCCCTGGCACCTCTTCCTTT- -AGTCTCCCGGCCCTACTTCATCTGAGTCACCCCATCA | 1908 |
| QY | 894 | ccatggggagggtgggctccaatctaaggcccttcctctgtcagaagggggttagggcaaaa | 953 |
| Db | 1909 | CCATGGAAGGTGGGGCCCTCAAACTAAGAACCTCTCTTTATGGAAGAAGGCGTGTGGCAAAA | 1968 |
| QY | 954 | gcccattacaagctgcctacccctcccccctttcagtgagccctgtgagccaggtgcttttc | 1013 |
| Db | 1969 | GCCCGCTATCAAACTGCCATTTCTTATGATTTCTGTGGACCTTGTGGCCAGGTGCTCTTC | 2028 |
| QY | 1014 | ctctatccacaggggtgtttgtgtgttggtgggtgtg | 1046 |
| Db | 2029 | TCGATCCACAGGTGTGTGTGAGCTTGTCTTTTG | 2061 |

Search completed: September 5, 2002, 16:33:29
Job time: 12132 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 18:30:02 ; Search time 457.59 Seconds
(without alignments)
4022.230 Million cell updates/sec

Title: US-09-824-647-16_COPY_1024_2095
Perform score: 1072
Sequence: 1 gggcccccaccagtgccctg.....ataagttgtgactttttt 1072

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-------------|--------------------|
| 1 | 1070.4 | 99.9 | 2095 | 20 AAV82825 | Human GP88 autocri |
| 2 | 1070.4 | 99.9 | 2124 | 14 AAQ56794 | Granulin coding se |
| 3 | 1070.4 | 99.9 | 2198 | 14 AAQ49052 | Granulin coding se |
| 4 | 992.6 | 92.6 | 2432 | 21 AAC78180 | Human cancer assoc |
| 5 | 989.4 | 92.3 | 1630 | 22 AAH16370 | Human cDNA sequenc |
| 6 | 747.2 | 69.7 | 1779 | 12 AAQ14339 | Human epithelin pr |
| 7 | 535.4 | 49.9 | 2137 | 20 AAQ62824 | Mouse GP88 autocri |
| 8 | 483.8 | 45.1 | 561 | 21 AAA69779 | Human ovarian carc |
| 9 | 469.2 | 43.8 | 1767 | 12 AAQ14338 | Rat epithelin prec |

| | | | | | | |
|----|-------|------|---------|----|-----------|--------------------|
| 10 | 448.4 | 41.8 | 1767 | 12 | AAQ14340 | Mouse epithelin pr |
| 11 | 339.2 | 31.6 | 439 | 16 | AAQ22115 | Human gene signatu |
| 12 | 228.8 | 19.3 | 233 | 22 | AAK18681 | Human brain expres |
| 13 | 206.6 | 19.3 | 565 | 22 | AAK53492 | Murine transport a |
| 14 | 197.8 | 18.5 | 336 | 22 | AAK54602 | Human ovarian PCR- |
| 15 | 197.8 | 18.5 | 336 | 22 | AAH83224 | Human ovarian tumo |
| 16 | 195.6 | 18.2 | 390 | 22 | AAK05895 | Human brain expres |
| 17 | 153.2 | 14.3 | 539 | 12 | AAQ14952 | Bovine epithelin p |
| 18 | 142.2 | 13.3 | 757 | 22 | AAH07141 | Human cDNA clone (|
| 19 | 65.4 | 6.1 | 430 | 22 | AAI82981 | Human polynucleoti |
| 20 | 64.6 | 6.0 | 341 | 12 | AAQ14953 | Chicken epithelin |
| 21 | 57 | 5.3 | 177 | 22 | AAI15045 | Oryctolagus cunicu |
| 22 | 51.8 | 4.8 | 561 | 21 | AAK69780 | Human ovarian carc |
| 23 | 43.8 | 4.1 | 561 | 21 | AAK69780 | Human ovarian carc |
| 24 | 43.4 | 4.0 | 1806 | 21 | AAK50102 | Human Znt2 consen |
| 25 | 43 | 4.0 | 2608 | 22 | AAK98385 | Human cDNA clone B |
| 26 | 38.2 | 3.6 | 199 | 14 | AAQ49051 | Probe (grnA) used |
| 27 | 38.2 | 3.6 | 114955 | 20 | AAK53491 | Human adenosine A1 |
| 28 | 38 | 3.5 | 774 | 17 | AAI13276 | DNA encoding uropo |
| 29 | 38 | 3.5 | 6741 | 21 | AAK10595 | Gene encoding a su |
| 30 | 37.8 | 3.5 | 10732 | 21 | AAI10594 | Gene encoding a su |
| 31 | 37 | 3.5 | 20303 | 18 | AAK71699 | Human deoxycytidyl |
| 32 | 37 | 3.5 | 26764 | 18 | AAK71696 | Human deoxycytidyl |
| 33 | 36.6 | 3.4 | 461 | 22 | AAK93033 | Human cDNA 3'-end |
| 34 | 36.6 | 3.4 | 1721 | 24 | ABK09743 | Human ovarian tumo |
| 35 | 36.6 | 3.4 | 2348 | 22 | AAH17671 | Human cDNA sequenc |
| 36 | 36.4 | 3.4 | 1501 | 22 | AAH24640 | Human cDNA sequenc |
| 37 | 36.4 | 3.4 | 2497 | 22 | AAH14473 | Human PEG1 DNA. B |
| 38 | 36.4 | 3.4 | 3789 | 23 | AAK86522 | DNA encoding novel |
| 39 | 36.4 | 3.4 | 18678 | 22 | AAK32889 | Human genomic DNA |
| 40 | 36.4 | 3.4 | 18678 | 22 | AAK91271 | Human digestive sy |
| 41 | 36.2 | 3.4 | 4403765 | 22 | AAI199683 | Mycobacterium tube |
| 42 | 36 | 3.4 | 935 | 21 | AAK29743 | Human gamma 1 heav |
| 43 | 36 | 3.4 | 935 | 21 | AAK29881 | Human gamma 1 heav |
| 44 | 36 | 3.4 | 935 | 22 | AAK30599 | Human gamma 1 heav |
| 45 | 36 | 3.4 | 1383 | 21 | AAK29742 | Human gamma 1 heav |

ALIGNMENTS

RESULT 1

| | |
|----------|---|
| AAV82825 | AAV82825 standard; cDNA; 2095 BP. |
| ID | AAV82825 standard; cDNA; 2095 BP. |
| XX | |
| AC | AAV82825; |
| XX | |
| DT | 15-MAR-1999 (first entry) |
| XX | |
| DE | Human GP88 autocrine growth factor cDNA. |
| XX | |
| KW | GP88; granulin; epithelin; human; growth factor; autocrine; tumour; |
| KW | cancer; viral infection; antagonist; therapy; diagnosis; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FH | Key |
| FT | CDS |
| FT | Location/Qualifiers |
| FT | 13..1794 |
| XX | /*tag= a |
| XX | |
| FN | W09852607-A1. |
| XX | |
| PD | 26-NOV-1998. |
| XX | |
| PF | 22-MAY-1998; 98WO-US10555. |
| XX | |
| PR | 16-DEC-1997; 97US-0991862. |
| PR | 23-MAY-1997; 97US-0863079. |
| XX | |
| PA | (SERRR/) SERRERO G. |
| XX | |
| PI | Serrero G; |
| XX | |

| | |
|----------|----------------------------------|
| RESULT | 6 |
| AAQ14339 | |
| ID | AAQ14339 standard; DNA; 1779 BP. |
| XX | |
| XX | AAQ14339; |
| XX | |
| DT | 17-JAN-1992 (first entry) |
| XX | |
| DE | Human epithelin precursor. |
| XX | |

KW XX EP; anti-sense RNA; growth regulation; inhibition; stimulation; ss.
 XX XX Homo sapiens.
 XX XX
 XX XX
 FH XX Location/Qualifiers
 FT misc_RNA 41..1819
 FT /*tag= a
 FT /note= "claim 9, page 54"
 XX WO91115510-A.
 XX 17-OCT-1991.
 XX PD
 XX 03-APR-1991; 91WO-US02321.
 XX PF
 XX 13-MAR-1991; 91US-0083796.
 XX PR 03-APR-1990; 90US-0504508.
 XX (BRIM) BRISTOL-MYERS SQUIB.
 XX PA
 XX Shoyab M, Plowman GD;
 XX PI
 XX WPI; 1991-325168/44.
 XX DR P-PSDB; AAR14326.
 XX XX
 XX New cysteine-rich growth modulating proteins, epithelins - useful
 PT as inhibitors of neoplastic cell growth and to promote wound
 PT healing and treat psoriasis
 PT
 XX Disclosure; Fig 22; 97pp; English.
 XX PS
 XX ET-1 and ET-2 were isolated from rat kidneys and their amino acid
 CC sequences determined. A full length rat ER cDNA (AAQ14338) was obt'd.
 CC by screening a rat kidney cDNA library in lambda gt10 with PCR
 CC generated ET probes. These probes were also used to obtain the mouse
 CC ET gene (AAQ14340) from a mouse T-cell genomic library.
 CC ET DNA was also obtained from human sources (AAQ14339).
 CC An anti-sense ribonucleic acid molecule complementary to the
 CC indicated fragment in the features is also claimed.
 CC ET-1 is a bifunctional growth regulator, capable of stimulating
 CC the growth of some cell types while inhibiting the growth of others.
 CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory
 CC bioactivity. In contrast, however, ET-2 is apparently not capable of
 CC eliciting the growth stimulatory activity characteristic of ET-1 and,
 CC in fact, antagonises this ET-1 activity.
 CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.
 XX CC
 XX Sequence 1779 BP; 304 A; 578 C; 547 G; 350 T; 0 other;
 XX SQ
 Query Match 69.7%; Score 747.2; DB 12; Length 1779;
 Best Local Similarity 98.3%; Pred. No. 1.5e-185;
 Matches 755; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 1 gggccaccaggctgcctggatggagaagcccccagctcacctcagctccagaccaca 60
 Db 1012 gggccaccaggctgcctggatggagaagcccccagctcacctcagctccagaccaca 1071
 QY 61 caagcctgaagagagatgtccctgtgataatgtcacagctgtccctcctccgatacc 120
 Db 1072 caagccttgaagagagatgtccctgtgataatgtcacagctgtccctcctccgatacc 1131
 QY 121 tgcctgccaactcacgctctggggagtggggctgtgtccaatccacagagctgtctgtcgc 180
 Db 1132 tgcctgccaactcacgctctggggagtggggctgtgtccaatccacagagctgtctgtcgc 1191
 QY 181 tcggaccaccagcactgtctccccagcgtacacagctgttagctgagggcagctgcag 240
 Db 1192 tcggaccaccagcactgtctccccaggggtcacacgctgtgtacgtgagggcagctgcag 1251
 QY 241 cgaaggagcagatcgtggctggactggagaagatgctgcgcccgccggttctctatcc 300
 Db 1252 caagagaacagatcgtgaactggactgaagagatgctgcgcccgagcttcttacc 1311

XX PD 17-OCT-1991.
XX PF 03-APR-1991; 91WO-US02321.
XX PR 13-MAR-1991; 91US-0083796.
XX PR 03-APR-1990; 90US-0504508.
XX PA (BRIM) BRISTOL-MYERS SQUIB.
XX PI Shoyab M, Plowman GD;
XX DR WPI; 1991-325168/44.
XX DR P-PSDB; AAR14325.
XX
PT New cysteine-rich growth modulating proteins, epithelins - useful
PT as inhibitors of neoplastic cell growth and to promote wound
PT healing and treat psoriasis
XX
PS Disclosure; Fig 18; 97pp; English.
XX
CC ET-1 and ET-2 were isolated from rat kidneys and their amino acid
CC sequences determined. A full length rat ET cDNA was obtd. by screening
CC a rat kidney cDNA library in lambda gt10 with PCR generated ET probes.
CC These probes were also used to obtain the mouse ET gene (AAQ14340) from
CC a mouse T-cell genomic library. ET DNA was also obtained from human
CC sources (AAQ14339).
CC An anti-sense ribonucleic acid molecule complementary to the
CC indicated fragment in the features is also claimed.
CC ET-1 is a bifunctional growth regulator, capable of stimulating
CC the growth of some cell types while inhibiting the growth of others.
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory
CC bioactivity. In contrast, however, ET-2 is apparently not capable of
CC eliciting the growth stimulatory activity characteristic of ET-1 and,
CC in fact, antagonises this ET-1 activity.
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.
XX
SQ Sequence 1767 BP; 359 A; 499 C; 491 G; 418 T; 0 other;

Query Match 43.8%; Score 469.2; DB 12; Length 1767;
Best Local Similarity 76.7%; Pred. No. 7.3e-113;
Matches 589; Conservative 0; Mismatches 173; Indels 6; Gaps 1;

QY 1 gggcccccacagtgccctggatgagagggccacgctccactcagctccagccagacc 60
DB 1006 ggaagcttcaggtaacctggatgagagggccacgctccactcagctccagccagacc 1065
QY 61 caagccttgagagagatgcccctgtgataatgtcagcagctgtccctccctccagacc 120
DB 1066 cagatcttgaagaatgatgcccctgtgactcagctcagtagctgtccttcaacaatcc 1125
QY 121 tcttcccaactcacgtctggagtgagggtggctgtctccaatccagagctgtctgtgc 180
DB 1126 tgcgtcagactcagctctgggagctgggctgtctccatcccaagagctgtctgtgc 1185
QY 181 tcggaccacacagctgtgcccacagcagacacagctgtgtagctgagggcagctgtcag 240
DB 1186 ttgacacacacagctgtgcccacagcagacacagctgtgtagctgagggctagctcag 1245
QY 241 caggaagcagagatgctggctggagctgaggaagatgctgcccgcgcgcgggttcctatcc 300
DB 1246 aaggagacagagaatggtggctggcctgaggaagatgctgcccgcgcgcgaactctgctc 1305
QY 301 caacccagagacatggctgtgacacagacacacacagctcccggtggcggaacctgtctgc 360
DB 1306 caacatgagagatgtgtgtgacacagacacacacagctcccggtggcggaacctgtctgc 1365
QY 361 ccagaccaggggtggagctgggctgtgcccagctgcccagctgtgtgtgtgtgtgtgtgt 420
DB 1366 ccaagcctgaagggaagttgggctgtgcccagctgcccagctgtgtgtgtgtgtgtgtgt 1425
QY 421 cgcacagcagctgtgcccggctgggtacacacctgcaacgtgaaggctgcatcctgcgagag 480

DB 1426 cggcagcactgttgcgggctgggtacacctgcaacgtgaaaggcagaaacctgtgagaag 1485
QY 481 gaagtgtctctgccacagcctgcccacaccttccctggcccgtagccctcactggtgtgag 540
DB 1486 gatcagcctctgtccagccttccatgacctgaccttggctcctaaggttgg-----g 1539
QY 541 gacgtgagtgagggaagacacttctgccaatgataacacacagcctgtgcccagagcaac 600
DB 1540 aatgtggaatgtggtgcgggacatttctgcccataacacacagcctgttctgtaagacagc 1599
QY 601 cgacaggctggtgctgtctccctacgcccagggcgctgtgtgtgtgtgtgtgtgtgtgtgt 660
DB 1600 caagaggctggtgctgtctccctatgttaaagggtgtgtgtgtgtgtgtgtgtgtgtgt 1659
QY 661 tgcgtcctgtggttccctgctgcagcaggggtaccaggtgtgtgtgtgtgtgtgtgtgtgtgt 720
DB 1660 tgtgtccattggttccactgttccagcgaagggaacacaggtgtgtgtgtgtgtgtgtgtgt 1719
QY 721 cgcgctggtgagcggccttggaggaccacacagcctgtgagacagctgtgtgtgtgtgtgt 768
DB 1720 cctcgtggacatacttggaggatccagcccaagaccgctactg 1767

RESULT 10
AAQ14340
ID AAQ14340 standard; DNA; 1767 BP.
AC AAQ14340;
XX 17-JAN-1992 (first entry)
DE Mouse epithelin precursor.
DE Mouse epithelin precursor.
KW ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.
XX Mus musculus.
OS Mus musculus.
FH Key 8..1774
FT misc_RNA
FT Location/Qualifiers
FT /*tag= a
FT /note= "Claim 29, page 56"
XX WO9115510-A.
XX 17-OCT-1991.
XX 03-APR-1991; 91WO-US02321.
XX 13-MAR-1991; 91US-0083796.
XX 03-APR-1990; 90US-0504508.
XX (BRIM) BRISTOL-MYERS SQUIB.
XX Shoyab M, Plowman GD;
XX WPI; 1991-325168/44.
XX P-PSDB; AAR14327.
XX New cysteine-rich growth modulating proteins, epithelins - useful
XX as inhibitors of neoplastic cell growth and to promote wound
XX healing and treat psoriasis
XX Disclosure; Fig 23; 97pp; English.
XX ET-1 and ET-2 were isolated from rat kidneys and their amino acid
XX sequences determined. A full length rat ET cDNA (AAQ14338) was obtd.
XX by screening a rat kidney cDNA library in lambda gt10 with PCR
XX generated ET probes. These probes were also used to obtain the mouse
XX ET gene (AAQ14340) from a mouse T-cell genomic library.
XX ET DNA was also obtained from human sources (AAQ14339).
XX An anti-sense ribonucleic acid molecule complementary to the
XX indicated fragment in the features is also claimed.

Db 241 catcacatggaggtgggcccataataagc-cttccctgtaaaggggtgtg 299
QY 949 caaaagcc-cattacaagctggccatccctcccggtttcagtgagaccctgtggccagtg 1007
Db 300 caaaagccattacaagctggccatccctcccggtttcagtgagaccctgtggccagtg 359
QY 1008 cttttccctatccacaggggtgttt 1032
Db 360 cttttccctatccacaggggtgttt 384

RESULT 12
AAK18681/c
ID AAK18681 standard; DNA; 233 BP.
XX
AC AAK18681;
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 18672.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PS WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human brains.
XX
PS Example 4; SEQ ID NO: 18672; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.
XX
SQ Sequence 233 BP; 43 A; 75 C; 73 G; 42 T; 0 other;

Query Match 21.3%; Score 228.8; DB 22; Length 233;
Best Local Similarity 99.18; Pred. No. 2.8e-50;
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 403 gctgtgtctgcgagatccgcagcactgtgccggtgtgtacacgtgcaactggaag 462
Db 233 GCTGTGTCTGCGAGATCCGCAGCAGCTGTGTCGCCGGTGTACACCTGCAACGTGAAG 174
QY 463 gctgatcctgcgagaagaagtgtctgtccagcgtgcaccttctctggtcccgtagc 522

Db 173 GCTGATCTCTCGAAGAGAGTGTCTCTGCCAGCCTGCCACCTTCTGCGCCGTAGC 114
QY 523 cctcacgtgggtgtgaaggacgtgagtggtgggaagacacttctgcatgataaccag 582
Db 113 CCTCACGTGGGTGTGAAGGACGTGGAGTGTGGGAAGGACACTTCTGCGCATGATAACCAG 54
QY 583 acctgtgcccagagacaaccacaggggtggcctgctgtcctcaccgcccagg 634
Db 53 ACTTGTCTGCCGAGACAACCGACAGGGCTGGGCTGCTGCTCCTACCGCCAGG 2

RESULT 13
AAK53492/c
ID AAK53492 standard; cDNA; 565 BP.
XX
AC AAK53492;
DT 16-NOV-2001 (first entry)
XX
DE Murine transport and binding associated protein encoding cDNA SEQ ID 57.
XX
KW Murine; liver; gene library; amino acid synthesis; binding protein;
KW cell metabolism; energy metabolism; fatty acid metabolism; synthesis;
KW phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;
KW replication; transcription; translation; transport protein; ss.
XX
OS Mus musculus.
XX
PN DE20103510-U1.
XX
PD 07-JUN-2001.
XX
PF 28-FEB-2001; 2001DE-2003510.
XX
PR 02-DEC-1999; 99DE-1058160.
XX
PA (LION-) LION BIOSCIENCE AG.
XX
DR WPI; 2001-368570/39.
XX
PT Gene library containing sequences with specific 3'-ends and no polyA tail, encoding proteins involved in a wide range of cellular processes
XX
PS Claim 15; Page 42; 251pp; German.
XX
CC This invention describes a novel gene library (A) comprises a gene sequence (or its part) encoding a protein involved in amino acid synthesis, cellular/energy metabolism, metabolism of fatty acids/phospholipids, synthesis or breakdown of purines/pyrimidines/nucleosides/nucleotides, DNA replication/transcription/translation, or is a transport/binding protein.
XX
CC (A) are produced that correspond to the 3'-end of mRNA but without the polyA tail. They can be prepared more efficiently and with less effort than conventional libraries. AAK53436-AAK54275 represent fragments of the gene library described in the method of the invention.
XX
SQ Sequence 565 BP; 144 A; 150 C; 151 G; 120 T; 0 other;

Query Match 19.3%; Score 206.6; DB 22; Length 565;
Best Local Similarity 68.0%; Pred. No. 2.4e-44;
Matches 349; Conservative 0; Mismatches 154; Indels 10; Gaps 4;

QY 544 gtgagtggtgggaaggacactctgtccatgataaccagcctgtgctgcg-agacaacg 602
Db 565 GTGAGTGTGGAAAGGGCAVTTCTGCATGATACACAGACCTGTGTAAAGACAGTGC 506
QY 603 acagggtggcctgctgtccctacgcccagggcgtctgtgtgtgtgctgctgctg 662
Db 505 AGGAGTCTGGGCTGCTGTCCTACCTAAAGGGTGTCTCTGTAGAGATGGAGCTCAGT 446

```
Qy 663 ctgtctctgcttgcctcgcgcagcaggggtaccagtggtttgagcaggggggcccc 722
Dy 445 TTGCCCGGCTGCTTCCACTGTTAGCCAGGGAACCAAGTGTGCGAAGAAAGATTCC 386
Qy 723 gcgctggagcgccttggagggaccagccttgagacagctgtgagggacagttact 782
Dy 385 TCGCTGGGACATGTTTTGAGGGATCCGCTCCCAAGACCGCTACTGTAAAGGAAGGGCTAC 326
Qy 783 gaa-----gactctgcagccctcgcgcagcccccagcctcggaggggtgcccctctgcagggc 835
Dy 325 AGACTTAAGGAACCTCCACAGTCTTGGGAGCCCTGTTCCGAGGGGTACCCACTACTCAGGCC 266
Qy 836 tccctagacacctccccttaacaaattctcctggagcccccattgtgag-ctcccatac 894
Dy 265 TCCCTAGGCGCTCTCCCTTAACGCTCTCCCGGCTACTCTCTGAGTACACCTATAC 206
Qy 895 catggaggtgggctcaattcaagcccttccctctcagagcctgtgcagaggggttgagg 953
Dy 205 CATGGAGGTGGAGGCTCAACTTAACCTTCTTTTATGGAAGAGGCTGTGGCCAAA 146
Qy 954 gccattacaagctgccatccccctccccttctcagtgagcctgtggccaggtgtctttc 1013
Dy 145 GCCCGTATCAAACTGCGCATTTCTCCGGTTTCTCTGGACCTTGTGGCCAGGTGCTCTTC 86
Qy 1014 cctatccacaggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1046
Dy 85 CCGAGCCACAGGTGTCTGTGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 53

RESULT 14
AAS24602/C
ID AAS24602 standard; cdna; 336 BP.
XX
AC AAS24602;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #783.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
OS Homo sapiens.
XX
PN WO200157207-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03733.
XX
PR 04-FEB-2000; 2000US-0180403.
PR 28-MAR-2000; 2000US-0192745.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Mannion J;
XX
XX WPI; 2001-488879/53.
XX
XX New polynucleotides encoding ovarian tumour proteins, useful for
XX treating ovarian cancer, and as probes, primers, and markers of cancer
XX progression -
XX
XX Example 1; page 237; 378pp; English.
XX
XX The invention comprises compositions used for the therapy and diagnosis
XX of ovarian cancer. The compositions comprise one or more ovarian tumour
XX proteins, their associated polynucleotides, or immunogenic portions of
XX the proteins. The ovarian tumour polynucleotides and polypeptides are
XX useful for stimulating and/or expanding T cells specific for a tumour
XX protein. They are also useful for inhibiting the development of cancer in
XX a patient with an ovarian tumour DNA or protein by incubating isolated
```

```
CC T-cells allowing them to proliferate, and administering to the patient..
CC The sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
CC represent human ovarian tumour protein CDNA clones.
XX
SQ Sequence 336 BP; 70 A; 87 C; 106 G; 62 T; 11 other;
Query Match 18.5%; Score 197.8; DB 22; Length 336;
Best Local Similarity 97.18; Pred. No. 4e-42;
Matches 233; Conservative 0; Mismatches 3; Indels 4; Gaps 3;
Qy 764 tgcgtgagggac-agtactgaagactctgcagcctcggagccctcggagccactcgaggggtgcc 822
Dy 239 TGCTGTGAGGACAAAGTACTGAAGACTTTGCAGCCCTCGGACCCCACTCGGAGGGTGCC 180
Qy 823 ctctgctcagcctcctcctagcacctccccttaacaaattctcctggagccactcttga 882
Dy 179 CTCGTNTCAGGCTCTCCCTAGCACCTCCCTTAACCAAAATTCCTCGGACCCCAATCTGA 120
Qy 883 gctcccatcaccatggaggtggggtcctcaatctaagcccttccctcctcagaagggg 942
Dy 119 GCTCCCCATCACCATGGGAGGTGGGGCTCAATCTAAGG-CTTCCCTGTGAGAGGGGG 61
Qy 943 ttgagggcaaaagcc--cattacaagctgccatccccctcccgtttcagtgagccctgtgg 1000
Dy 60 TTGTGGCAAAAGCCACATTACAGCTGCCATCCCTCCCGGTTTCAGTGGACCTGTGG 1

RESULT 15
AAH83224/C
ID AAH83224 standard; cdna; 336 BP.
XX
AC AAH83224;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:848.
XX
KW Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
KW immunogenic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200151513-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01575.
XX
PR 14-JAN-2000; 2000US-0176722.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA;
XX
XX WPI; 2001-425866/45.
XX
XX Novel ovarian tumor proteins, and nucleic acids encoding them, used to
XX treat and diagnose cancers, particularly ovarian cancer -
XX
XX Claim 5; Page 219; 338pp; English.
XX
XX AAH82377 to AAH83878 represent human ovarian tumour-associated
XX polynucleotide sequences which encode ovarian tumour proteins. The
XX ovarian tumour proteins and polynucleotide sequences have cytostatic
XX activity, and can be used in gene therapy and vaccine production. The
XX ovarian tumour proteins and polynucleotides can be used to inhibit
XX the development of cancer, particularly ovarian cancer. They can also
XX be used to diagnose the onset and progression of cancer.
XX
```

OS
XX

SQ Sequence 336 BP; 70 A; 87 C; 106 G; 62 T; 11 other;

Query Match 18.5%; Score 197.8; DB 22; Length 336;

Best Local Similarity 97.18; Pred. No. 4e-42;

Matches: 233; Conservative: 0; Mismatches: 3; Indels: 4; Gaps: 3;

Qy 764 **tgctgtgaggac-agtactgaagactctgcagccctcgggaccccaactcgagggtgcc** 822

[illegible]

QY 823 ctctgctcaggcctcccttagcacctccccctaaccaaatctccctggacccattctga 882

—

Qy 883 gctcccatcaccatgggaggtggggcctcaatctaaggcccttccctgtcagaagggg 942

[illegible]

Qy 943 ttgaggcaaaagcc--cattacaagctgccatccccctccccgttttcagtggaacctgtgg 1000

[illegible]

Search completed: September 5, 2002, 18:30:15
Job time: 10547 sec

Job time: 10547 sec